Moment estimation in discrete shifting level model applied to fast array-CGH segmentation

Statistica Neerlandica (2013) Vol. 67, nr. 3, pp. 227–262
doi:10.1111/stan.12005

Moment estimation in discrete shifting level model applied to fast array-CGH segmentation

A. Gandolfi*

Dipartimento di Matematica U. Dini, Università di Firenze, Viale Morgagni 67/A, 50134 Florence, Italy

M. Benelli

Dipartimento di Matematica U. Dini, Università di Firenze, Viale Morgagni 67/A, 50134 Florence, Italy and Center for the Study of Complex Dynamics (CSDC), University of Florence, Florence, Italy

A. Magi

Dipartimento di Matematica U. Dini, Università di Firenze, Viale Morgagni 67/A, 50134 Florence, Italy and Diagnostic Genetic Unit, Careggi University Hospital, 50134 Florence, Italy

S. Chiti

Dipartimento di Matematica U. Dini, Università di Firenze, Viale Morgagni 67/A, 50134 Florence, Italy

We develop a mathematical theory needed for moment estimation of the parameters in a general shifting level process (SLP) treating, in particular, the finite state space case geometric finite normal (GFN) SLP. For the SLP, we give expressions for the moment estimators together with asymptotic (co)variances, following, completing, and correcting CLINE (*Journal of Applied Probability* 20, 1983, 322–337); formulae are then made more explicit for the GFN-SLP. To illustrate the potential uses, we then apply the moment estimation method to a GFN-SLP model of array comparative genomic hybridization data. We obtain encouraging results in the sense that a segmentation based on the estimated parameters turns out to be faster than with other currently available methods, while being comparable in terms of sensitivity and specificity.

Keywords and Phrases: shifting level process, moment estimator, array-CGH, finite state space, segmentation, confidence intervals, DNA, microarray.

1 Introduction

We develop here a mathematical theory related to moment estimations of the parameters in a shifting level process (SLP) or shifting level model (Chernoff and Zacks,

*gandolfi@math.unifi.it

1964; Salas and Boes, 1980). In the a mathematical paper on such moment estimation, CLINE (1983) considered a general SLP $\mathbb{Y} = \{Y_\alpha\}_{\alpha=1}^{\infty}$, constructed as a concatenation of segments of random length, randomly selected from a family of processes (all of the mechanisms describing such randomness being identified as the underlying processes), and derives, under very general conditions, asymptotic properties of the empirical moments

$$\frac{1}{\alpha} S_\alpha^f = \frac{1}{\alpha} \sum_{i=1}^{\alpha} f(Y_i, Y_{i+1} \ldots, Y_{i+k})$$

with $f : \mathbf{X}^{k+1} \to \mathbb{R}^P$. In particular, CLINE (1983) managed to derive, under suitable but very general conditions, a law of large numbers and a central limit theorem (CLT) for $\frac{1}{\alpha} S_\alpha^f$ as functions of the moments of the underlying processes. In Section 2 and Appendix A, we recall Cline's main results, obtaining then more explicit and readable formulae when f is a polynomial (which amounts to all what is needed in our intended main application) and correcting two mistakes in Cline's paper (Appendix B).

CLINE (1983) then specialized to an SLP with geometrically distributed segment lengths and other underlying processes being normal (a geometric normal normal or GNN-SLP), and provides, without showing the very long calculations, explicit formulae for asymptotic moments and their (co)variances.

Here, instead, we specialize in a different direction, namely to a geometric finite normal shifting level process or GFN-SLP, in which segment lengths are still geometrically distributed and errors are normally distributed, but the state space is finite. For such case, we obtain more explicit formulae in Section 3 for the asymptotics of the empirical moments. Detailed calculations are demonstrated in the Appendix, where we also correct two errors in Cline's paper. In particular, we manage to invert the asymptotic expressions in Lemma 2 for the first moments and 2-autocovariances and 3-autocovariances as functions of the model parameters. This allows to explicitly determine moment estimators and their asymptotic (co)variances. These are the main results of this paper.

To illustrate the potential applications of the moment estimations in the GFN-SLP, we consider the segmentation problem in array comparative genomic hybridization (array-CGH) data.

Array-CGH (OOSTLANDER, MEIJER and YLSTRA (2004)) is a microarray technique that allows detection and mapping of genomic alterations (see CARTER (2007)). Test and reference DNA are differently fluorescent labeled, arrays of clones are accurately spotted (following human genome) onto glass slides, and then the mixed fluorescent DNA is hybridized to the array. The resulting fluorescent ratio is then measured, clone by clone, with measurements affected by a non-negligible noise; currently, one array can contain up to 10^6 probes, each of the order of 20–100 monomers (LIU (2007)). A function (the log base 2) of the fluorescent ratio is then plotted as function of the clone number, giving a discrete time jump process. In the

subsequent array-CGH analysis, one needs to detect the breakpoints where there is DNA copy-number variations (CNVs) and then identify for each connected region the copy number, calling *neutral* for the physiological two copies, *loss* for less, and *gain* for more copies. The task is complicated by the high level of noise in the measurement process, which confuses short segments with CNV with a noisy but physiologically normal tract.

After several segmentation methods have been devised (HUPE *et al*. 2004; PICARD *et al*. 2005; OLSHEN *et al*. 2004; MYERS *et al.* 2004), in MAGI *et al*. (2010), the GNN version of the SLP (GNN-SLP) has been successfully used to model and analyze array-CGH data. In the approach of MAGI *et al*. (2010), array-CGH data are modeled by a GNN-SLP, and the analysis consists of assigning a preliminary segmentation and then carrying out an iterative approach similar to the pseudo-expectation–maximization algorithm for hidden Markov models (HMMs) FORTIN and KEHAGIAS (2006): a partly iterative estimation of number of states and model parameters (FORNEY (1973)) is performed in the E step and, finally, the best segmentation is obtained in the M step by using the Viterbi algorithm. The E and M steps are repeated until an identical result is obtained. The algorithm is approximately quadratic in the number of probes, and, although it is not yet the case, this might turn out to be a critical issue as the number of probes is dramatically increasing with technological advances.

We follow here a similar approach, which is presented in Section 5. However, we start in Section 4 by noticing that the state space of the SLP is not arbitrary, as it reflects the possible values of (the log of) the fluorescent ratio of DNA copy number against normal; such ratio can only be $0,1/2,1,3/2,\ldots$, with some noise due to the color reading mechanism, and occasional minor alterations due to genetic reasons. Notice that level 1 reflects normality. By these remarks, the state space of the SLP contains only few rather well-determined values, which can be separately determined at the start of the analysis, possibly using previous genetic information; to avoid missing unusual values, it is also possible to include extra states (as long as this does not burden running time, this has no lasting effect as probabilities estimations permit to identify irrelevant states). We are then modeling the array-CGH data as an SLP with geometric waiting time (G) between switches, a finite distribution (F) over the previously identified states, and a normal independent noise (N) with constant variance. This amounts to a GFN-SLP, as described in Section 4.

In Section 5, we describe how to apply our method to the segmentation problem in array-CGH. Starting from the fixed set of possible states, we obtain the model parameters by using the moment estimators, and then we can apply just one step of the Viterbi algorithm to obtain a segmentation. The detailed theory of moment estimation, which we develop in this paper, would allow also to determine confidence intervals (CI): we only give one example in Section 5, as the evaluation of the error in asymptotic approximation requires more investigation.

We then report the results of systematic comparisons of the segmentation based on moment estimation with some other currently used segmentation methods. Tests are

performed using synthetic chromosomes generated by LAI *et al*. (2005). In Section 6, we compare the receiver operating characteristic (ROC) curves generated by our method (with several different choices for the initial state) with those of other methods, and find that they are comparable.

In Section 7, we compare the execution times, revealing that the moment segmentation is faster than the other methods.

The results on the proposed method are thus extremely encouraging, in particular because the rapid growth of microarray size and resolution requires segmentation algorithms with high computational performance.

An additional issue, raised by an anonymous reviewer of the manuscript, concerns the normality assumption for the noise. In this work, the normality is assumed as it appears to be a good approximation for normalized read counts data; see YOON *et al*. (2009); on the other hand, it is conceivable that other distributions could be more adequate. As the crucial mathematical step in our procedure is Lemma 2, which shows that the map from parameters to statistics is continuously invertible, it would be interesting to find general conditions for the noise distribution under which such invertibility is ensured.

2 Results for general SLP

In this section, we recall the definition of SLP together with some results from CLINE (1983); we then write some general expressions of useful moments.

Let $(\mathbf{X}, \mathcal{X})$, $(\Lambda, \mathcal{L})$, and $(\mathbf{N}, \mathcal{N})$ be measurable spaces, with $\mathbf{N} = \{1, 2, 3 \ldots\}$, and let $(\Omega, \mathcal{F}, \mathrm{P})$ be the underlying probability space.

Definition 1. *If* $\left\{ \left\{ X_j^{(\lambda)} \right\}_{j=1}^{\infty}, \lambda \in \Lambda \right\}$ is a family of stochastic processes on $(\Omega, \mathcal{F}, \mathrm{P})$ with elements in $\mathbf{X}$, and $\{N_n, \Lambda_n\}_{n=1}^{\infty}$ is a stochastic process in $(\Omega, \mathcal{F}, \mathrm{P})$ with elements in $\mathbf{N} \times \Lambda$, then the process

$$\begin{aligned} \{Y_\alpha\}_{\alpha=1}^{\infty} &= \left\{ X_1^{(\Lambda_1)}, X_2^{(\Lambda_1)}, \ldots, X_{N_1}^{(\Lambda_1)}, X_1^{(\Lambda_2)}, \ldots, X_{N_2}^{(\Lambda_2)}, X_1^{(\Lambda_3)}, \ldots \right\} \\ &= \left\{ \left\{ X_j^{(\Lambda_n)} \right\}_{j=1}^{N_n} \right\}_{n=1}^{\infty} \end{aligned} \tag{1}$$

is called a Shifting Level Process or SLP with epochs 'shift' $\{T_n\}_{n=1}^{\infty} = \{N_1 + \ldots + N_n\}_{n=1}^{\infty}$, levels $\{\Lambda_n\}_{n=1}^{\infty}$, and underlying process $\left\{ X_j^{(\lambda)} \right\}_{j=1}^{\infty}, \lambda \in \Lambda$.

See CLINE (1983) for comments on the definition. The SLP generally depends on the parameters in the distributions of the X_j's and $\{N_n, \Lambda_n\}_{n=1}^{\infty}$, which can be estimated through the observable process $\{Y_\alpha\}_{\alpha=1}^{\infty}$. Notice that for all $\alpha \in \mathbb{N}$, the random variable Y_α takes value in $\mathbf{X}$, so that if $f : \mathbf{X}^{k+1} \to \mathbb{R}$, the sample moments are

$$\frac{1}{\alpha} S_\alpha^f = \frac{1}{\alpha} \sum_{i=1}^{\alpha} f(Y_i, Y_{i+1}, \ldots, Y_{i+k}). \tag{2}$$

As mentioned in CLINE (1983), we only consider real-valued sample moments as the results are easily extended to vector-valued or continuous functions of the sample moments. The main estimation results will be expressed in terms of the auxiliary, unobservable random variables

$$R_n^f = \sum_{i=T_{n-1}+1}^{T_n} f(Y_i, \ldots, Y_{i+k}) = \sum_{j=1}^{N_n} f\left(X_j^{(\Lambda_n)}, X_{I_{j+1}}^{(L_{j+1})}, \ldots, X_{I_{j+k}}^{(L_{j+k})}\right)$$

where

$$L_j = \Lambda_{m_j}$$
$$I_j = j - \left(T_{m_j} - T_n\right)$$

and m_j satisfies

$$\left(T_{m_j} - T_n\right) < j \leq \left(T_{m_{j+1}} - T_n\right).$$

For instance, if $f : \mathbf{X}^3 \to \mathbb{R}$, then

$$R_n^f = \sum_{j=1}^{N_n-2} f\left(X_j^{(\Lambda_n)}, X_{j+1}^{(\Lambda_n)}, X_{j+2}^{(\Lambda_n)}\right) + f\left(X_{N_{n-1}}^{(\Lambda_n)}, X_{N_n}^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}\right) + f\left(X_{N_n}^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_{I_{N_n+2}}^{(L_{N_n+2})}\right). \tag{3}$$

For later convenience, we indicate

$$f_j^{(n)} = f\left(X_j^{(\Lambda_n)}, X_{I_{j+1}}^{(L_{j+1})}, \ldots, X_{I_{j+k}}^{(L_{j+k})}\right)$$

and

$$U_n^f = \sum_{j=1}^{n} R_j^f.$$

CLINE (1983) presented some general sufficient conditions for the law of large numbers and the CLT for $\frac{1}{\alpha}S_\alpha^f$. We recall here Corollaries 2.1 and 3.1 only, as they are enough to deal with the discrete version used in the applications discussed later. It is these results that are used by Cline in the second part of his paper, where there are some errors corrected in Appendix B:

Proposition 1. *(Corollary 2.3 in CLINE (1983)). Let* $\{Y_\alpha\} = \left\{\left\{X_j^{(\Lambda_n)}\right\}_{j=1}^{N_n}\right\}$ *be an SLP such that*

1. $\{N_n, \Lambda_n\}$ *is a sequence of random elements in* $\mathbf{N} \times \Lambda$ *with* $P[\Lambda_n = \Lambda_m] = 0, n \neq m.$

2. $\left\{X_j^{(\lambda)}\right\}$, $\lambda \in \Lambda$ *is a family of independent stochastic processes and independent of* $\{N_n, \Lambda_n\}$.

Let $f : \mathbf{X}^{k+1} \to \mathbb{R}$ *and define* R_n^f *and* S_α^f *as before.*

1. *If* $\{N_n, \Lambda_n\}$ *is stationary, ergodic and* $E[N_n] = \eta < \infty$, $E\left[R_n^f\right] = \eta\theta < \infty$, *then* $(1/\alpha)S_\alpha^f \to \theta$ *a.s.*
2. *If* $\{N_n, \Lambda_n\}$ *is l-dependent and* $E[N_n] \to \eta$, $E\left[R_n^f\right] \to \eta\theta$, $E\left[R_n^{|f|}\right] \to \eta\zeta$, *and* $V\left[R_n^f\right] \le Kn^\beta$, $V\left[R_n^{|f|}\right] \le Kn^\beta$, $V[N_n] \le Kn^\beta$, $\beta < 1$, *then* $(1/\alpha)S_\alpha^f \to \theta$ *a.s.*

Proposition 2. *(Corollary 3.1 in* CLINE *(*1983*)). Let* $\{Y_\alpha\} = \left\{\left\{X_j^{(\Lambda_n)}\right\}_{j=1}^{N_n}\right\}$ *be an SLP and* $f : \mathbf{X}^{k+1} \to \mathbb{R}$ *be such that*:

1. $\{N_n, \Lambda_n\}$ *is a strictly stationary,* ϑ*-mixing of random elements in* $\mathbf{N} \times \Lambda$ *with* $P[\Lambda_n = \Lambda_m] = 0$ *for* $n \neq m$ *and* $\sum_{j=1}^{\infty} \vartheta(j)^{1/2} < \infty$.
2. $\left\{X_j^{(\lambda)}\right\}$, $\lambda \in \Lambda$ *is a family of independent stochastic processes and independent of* $\{N_n, \Lambda_n\}$
3. $V\left[R_n^f\right] < \infty$, $V[N_n] < \infty$, $V\left[R_n^{|f-\theta|}\right] < \infty$.

If

$$\begin{aligned} \eta &= E[N_n], \\ \eta\theta &= E\left[R_n^f\right], \\ \eta\chi_j &= Cov\left[R_n^f - \theta N_n, R_{n+j}^f - \theta N_{n+j}\right], \ j \ge 0 \end{aligned}$$

then

$$\sqrt{\alpha}\left(\frac{1}{\alpha}S_\alpha^f - \theta\right) \to N\left(0, \gamma^2\right) \text{ in distribution}, \tag{4}$$

where

$$\gamma^2 = \chi_0 + 2\sum_{j=1}^{\infty}\chi_j. \tag{5}$$

The preceding two results express the asymptotic values of the sample moments in terms of the moments of N_n and R_n^f. In turn, CLINE (1983) provided in Section 4 some formulae without derivation for the moments of R_n^f in terms of the moments of N_n and $f\left(X_1^{(\Lambda_n)}, \ldots\right)$, provided that enough moments of f exist, but these are not directly computable in explicit examples as the moments of f require some careful direct computation depending on their different arguments. We give here some more

explicit and directly computable formulae for the moments of R_n^f, for polynomial f, in terms of the moments of N_n and the join moments of the $X_i^{(\Lambda_n)}$'s. This makes the derivation and verification of explicit expression much easier.

We now derive various formulae under the following hypothesis:

1. $\left\{X_j^{(\lambda)}\right\}$, $\lambda \in \Lambda$, is a family of independent stochastic processes, each of which is a sequence of exchangeable random elements of $\mathbf{X}$.
2. $\{N_n\}$ and $\{\Lambda_n\}$ are sequences of i.i.d. random elements of $\mathbf{N}$ and Λ, respectively, and are independent of each other and of $\left\{X_j^{(\lambda)}\right\}$.

The limit theorems require computing the moments of R_n^f. We compute them for $f((x_1 \ldots, x_r)) = \prod_{l=1}^r x_l^{h_l}$ in terms of the moments of N_n and of those of $\left\{X_j^{(\lambda)}\right\}$:

1. $\alpha_i = E\left[N_n^i\right]$,
2. $\beta_i(\lambda) = E\left[\left(X_j^{(\lambda)}\right)^i\right]$, and
3. $\mu_i = E[\lambda^i]$;

the last expression is not used in the first results below. Note that $\beta_i(\lambda)$ are random variables, and actually, the formulae are functions of the expected values of products of the $\beta_i(\lambda)$'s. Later, when we consider $\left\{X_j^{(\lambda)}\right\}$ to be normally distributed, we can substitute such expected values by formulae depending only on the moments of $\{\Lambda_n\}$.

We start with $f : \mathbf{X} \to \mathbb{R}$, that is, $f(x) = x^h$. At the price of additional complications in the formulae, we could deal with any analytic f, but we avoid such details here as they are not needed in the main applications below. Let

$$S_{k,r} = \{(k_1, \ldots, k_r, s_1, \ldots, s_r) : k_i \in \mathbb{N}, s_i \in \mathbb{N}, 1 \le k_1 < k_2 < \ldots < k_r \le k, \sum_{i=1}^r k_i s_i = k\}.$$

and

$$S_{k,r,N_n} = \{(k_1, \ldots, k_r, s_1, \ldots, s_r) \in S_{k,r} : s = \sum_{i=1}^r s_i \le N_n\}.$$

Then we have the following:

Theorem 1.

$$E\left[\left(R_n^{x^h}\right)^k\right] = \sum_{r=1}^k \sum_{(k_1,\ldots,k_r,s_1,\ldots,s_r) \in S_{k,r,N_n}} \frac{k!}{((k_1!)^{s_1} \cdot \ldots \cdot (k_r!)^{s_r})(s_1! \cdot \ldots \cdot s_r!)} \cdot E\left[\beta_{k_1 h}^{s_1} \cdot \ldots \cdot \beta_{k_r h}^{s_r}\right] \left(\alpha_s + \sum_{m=1}^{s-1} \alpha_{s-m} (-1)^m \sum_{1 \le i_1 < \ldots < i_m \le s-1} \prod_{j=1}^m i_j\right).$$

Proof

$$
\begin{aligned}
E\left[\left(R_n^{x^h}\right)^k\right] &= E\left[\left(\sum_{i=1}^{N_n}\left(X_i^{(\lambda)}\right)^h\right)^k\right] \\
&= E\left[\sum_{r=1}^{k}\sum_{(k_1,\ldots,k_r,s_1,\ldots,s_r)\in S_{k,r}}\frac{k!}{((k_1!)^{s_1}\cdot\ldots\cdot(k_r!)^{s_r})(s_1!\cdot\ldots\cdot s_r!)}\right. \\
&\quad \sum_{i_1,\ldots,i_s\in\{1,\ldots,N_n\},\ diffferent}\left(X_i^{1(\lambda)}\right)^{hk_1}\cdot\ldots\cdot\left(X_{i_{s_1}}^{(\lambda)}\right)^{hk_1}\left(X_{i_{s_1+1}}^{(\lambda)}\right)^{hk_2} \\
&\quad \left.\cdot\ldots\cdot\left(X_{i_{s_1+s_2}}^{(\lambda)}\right)^{hk_2}\cdot\ldots\cdot\left(X_{i_{s_1+s_2+\ldots s_{r-1}+1}}^{(\lambda)}\right)^{hk_r}\cdot\ldots\cdot\left(X_{i_s}^{(\lambda)}\right)^{hk_r}\right] \\
&= \sum_{r=1}^{k}\sum_{(k_1,\ldots,k_r,s_1,\ldots,s_r)\in S_{k,r}}\frac{k!}{((k_1!)^{s_1}\cdot\ldots\cdot(k_r!)^{s_r})(s_1!\cdot\ldots\cdot s_r!)} \\
&\quad \cdot E\left[\beta_{k_1h}^{s_1}\cdot\ldots\cdot\beta_{k_rh}^{s_r}\right]E[N_n(N_n-1)\cdot\ldots\cdot(N_n-s+1)] \\
&= \sum_{r=1}^{k}\sum_{(k_1,\ldots,k_r,s_1,\ldots,s_r)\in S_{k,r,N_n}}\frac{k!}{((k_1!)^{s_1}\cdot\ldots\cdot(k_r!)^{s_r})(s_1!\cdot\ldots\cdot s_r!)} \\
&\quad \cdot E\left[\beta_{k_1h}^{s_1}\cdot\ldots\cdot\beta_{k_rh}^{s_r}\right]\left(\alpha_s+\sum_{m=1}^{s-1}\alpha_{s-m}(-1)^m\sum_{1\le i_1<\ldots<i_m\le s-1}\prod_{j=1}^{m}i_j\right)
\end{aligned}
$$

where the third equality holds as the X_i^{λ}'s are conditionally independent given λ, the number of $i_1,\ldots,i_s\in\{1,\ldots,N_n\}$, different from each other, is $N_n(N_n-1)\cdot\ldots\cdot(N_n-s+1)$, and the variables N_n's are also independent. The last equality holds as $N_n(N_n-1)\cdot\ldots\cdot(N_n-s+1)=0$ for $s>N_n$. □

For the GNN model considered in Cline (1983), α_i is the ith moment of a geometric random variable with parameter π, $\beta_i(\lambda)$ is the ith moment of a $N(\lambda,(1-\rho)\sigma^2)$ distribution, and λ is itself $N(\mu,\rho\sigma^2)$ with ith moment μ_i. The four parameters of the model, ρ, σ, μ, and π, satisfy

$$
\begin{cases}
\rho\sigma^2=\mu_2-\mu_1^2\\
(1-\rho)\sigma^2=E[\beta_2(\lambda)]-(E[\beta_1(\lambda)])^2\\
\mu=\mu_1\\
\pi=\alpha^{-1}
\end{cases}
$$

from which we obtain their expression in terms of the moments β_1, β_2, μ_1, μ_2, and α:

$$
\begin{cases}
\sigma^2=E[\beta_2(\lambda)]-(E[\beta_1(\lambda)])^2+\mu_2-\mu_1^2\\
\rho=\dfrac{\mu_2-\mu_1^2}{E[\beta_2(\lambda)]-(E[\beta_1(\lambda)])^2+\mu_2-\mu_1^2}\\
\mu=\mu_1\\
\pi=\alpha^{-1}
\end{cases}.
$$

Then with the preceding formulae, it is easy to recompute (see Cline (1983) p. 334) for the function $f(x)=x$

$$
\begin{aligned}
E\left[R_n^x\right] &= E[\beta_1(\lambda)]\alpha_1 = E[\lambda]\frac{1}{\pi} = \frac{1}{\pi}\mu \\
V\left[R_n^x - \mu N_n\right] &= E\left[\left(R_n^x - \mu N_n\right)^2\right] = E\left[\left(R_n^x\right)^2\right] - 2\mu E\left[N_n R_n^x\right] + \mu^2 E\left[\left(N_n\right)^2\right] \\
&= E[(\beta_2(\lambda)]\alpha_1 + E\left[\left(\beta_1(\lambda)\right)^2\right](\alpha_2 - \alpha_1) - 2\mu\alpha_2 E[\beta_1(\lambda)] + \mu^2\alpha_2 \\
&= E\left[\lambda^2 + (1-\rho)\sigma^2\right]\alpha_1 + E\left[\lambda^2\right](\alpha_2 - \alpha_1) - 2\mu\alpha_2 E[\lambda] + \mu^2\alpha_2 \\
&= (1-\rho)\frac{\sigma^2}{\pi} + \left(\mu^2 + \rho\sigma^2\right)\frac{2-\pi}{\pi^2} - \mu^2\frac{2-\pi}{\pi^2} = \frac{1}{\pi}\sigma^2 + 2\frac{1-\pi}{\pi^2}\rho\sigma^2;
\end{aligned}
$$

and for $f(x) = x^2$ (denoted by f_0 in CLINE (1983))

$$
\begin{aligned}
E\left[R_n^{x^2}\right] &= \frac{1}{\pi}\left(\sigma^2 + \mu^2\right) \\
V\left[R_n^{x^2} - \left(\sigma^2 + \mu^2\right)N_n\right] &= E\left[\left(R_n^{x^2}\right)^2\right] - 2\left(\sigma^2 + \mu^2\right)E\left[N_n R_n^{x^2}\right] + \left(\sigma^2 + \mu^2\right)^2 E\left[\left(N_n\right)^2\right] \\
&= \left(\mu^2 + (1-\rho)\sigma^2\right)^2\alpha_1^2 - 2\left(\sigma^2 + \mu^2\right)\alpha_2\left(\mu^2 + (1-\rho)\sigma^2\right) + \alpha_2\left(\sigma^2 + \mu^2\right)^2 \\
&= \frac{1}{\pi}\left(2\sigma^4 + 4\mu^2\sigma^2\right) + 2\frac{1-\pi}{\pi^2}\left(2\rho^2\sigma^4 + 4\rho\mu^2\sigma^2\right).
\end{aligned}
$$

3 The GFN-SLP model

We now specialize to another particular SLP: the geometric finite normal or GFN-SLP. In the GFN-SLP, $\{N_n\} \sim$ i.i.d. geometric (π); given λ, $\left\{X_n^{\{\lambda\}}\right\} \sim$ i.i.d. $N(\lambda,\tau^2)$; and $\{\Lambda_n\} \sim$ i.i.d. with a finite distribution on $\{b_1,\ldots,b_T\}$ with parameters $\mathbf{p} = \{p_1,\ldots,p_T\}$; all processes being independent. To avoid trivialities and simplify the later formulae, we assume $T > 1$ and $\pi < 1$, which is to say, that $\{Y_i\}$ is not independent. The GFN-SLP is simple enough that Propositions 1 and 2 apply. In particular, α_i's, β_i's, and μ_i's can be explicitly written in terms of the parameters of the model. In fact, we have $\alpha_1 = 1/\pi$, $\alpha_2 = \frac{2-\pi}{\pi^2}$,

$$
\beta_i(\lambda) = \sum_{j=0}^{\lfloor i/2 \rfloor} \rho_j(i)\lambda^{i-2j}\tau^{2j},
$$

the ith moment of a normal (λ,τ^2) distribution, for which there exist explicit expressions for the $\rho_j(i)$; and, finally, $\mu_i = \mu_i(p_1,\ldots,p_T)$ is the ith moment of the finite distribution of the $\{\Lambda_n\}$'s.

We now intend to estimate the $T+1$ parameters $p_1,\ldots,p_{T-1}$, π and τ^2; for convenience, we consider the $T+2$ parameters $\mathbf{p}$, π, and τ^2 subject to the constraint $\sum_i p_i = 1$. The statistics used will be sample moments of the form given in (2) for the functions $f(x) = x^i$, $f_1(x_1,x_2) = x_1 \cdot x_2$, and $f_2(x_1,x_2,x_3) = x_1 \cdot x_2 \cdot x_3$. More precisely, we use

$$\begin{aligned}\hat{m}_i &= \frac{1}{n} S_n^{x^i}, \qquad i = 1, \ldots, T-1 \\ \hat{m}_{f_1} &= \frac{1}{n} S_n^{x_1 \cdot x_2} \\ \hat{m}_{f_2} &= \frac{1}{n} S_n^{x_1 \cdot x_3}.\end{aligned} \tag{6}$$

The next lemma computes the asymptotics of the statistics in terms of the model parameters, together with the asymptotic variances. Then the subsequent lemma shows how to explicitly invert such asymptotics to retrieve the model parameters; the final theorem gives the explicit form of the parameter estimators with their asymptotic variance.

Lemma 1. *If* $\{Y_\alpha\} = \left\{ \left\{ X_j^{(\Lambda_n)} \right\}_{j=1}^{N_n} \right\}$ *is a GFN-SLM, then for all* $i = 1, \ldots, T-1$

$$\hat{m}_i = \frac{1}{n} S_n^{x^i} \to m_i := \sum_{j=0}^{\lfloor i/2 \rfloor} \rho_j(i) \mu_{i-2j} \tau^{2j} \text{ a.s.} \tag{7}$$

and

$$\sqrt{n}\left(\frac{1}{n} S_n^{x^i} - m_i \right) \to N\left(0, \gamma_i^2\right) \textit{ in distribution},$$

where

$$\gamma_i^2 = \frac{1}{\alpha_1} \left[E\left[\left(R_n^{x^i} \right)^2 \right] - \alpha_2 m_i^2 \right] = \frac{1}{\alpha_1} \left[(\alpha_2 - \alpha_1) E\left[\beta_i^2(\lambda) \right] + \alpha_1 m_{2i} - \alpha_2 m_i^2 \right].$$

Moreover,

$$\hat{m}_{f_1} = \frac{1}{n} S_n^{f_1} \to m_{f_1} := \frac{1}{\alpha_1} \left[(\alpha_1 - 1)\mu_2 + \mu_1^2 \right] = \frac{1}{\alpha_1} \left[(\alpha_1 - 1)\left(m_2 - \tau^2 \right) + m_1^2 \right] \text{ a.s.}$$

and

$$\sqrt{n}\left(\frac{1}{n} S_n^{f_1} - m_{f_1} \right) \to N\left(0, \gamma_{f_1}^2\right) \textit{ in distribution},$$

where

$$\begin{aligned}\gamma_{f_1}^2 = \frac{1}{\alpha_1} \Bigg\{ & (\alpha_2 - 2\alpha_1 + 1) m_4 + \left(2\alpha_2 - 7\alpha_1 - 2\pi + 6 + \frac{\alpha_2}{\alpha_1^2} \right) \tau^4 \\ & + \left(2\alpha_1 - \alpha_2 - 1 + \frac{\alpha_2}{\alpha_1^2} \right) m_2^2 + \left(12\alpha_1 - 4\alpha_2 + 2\pi - 8 - 2\frac{\alpha_2}{\alpha_1^2} \right) m_2 \tau^2 \\ & + \left(\frac{\alpha_2}{\alpha_1^2} - 4 \right) m_1^4 + 4(\alpha_1 - 1) m_1 m_3 + 2\left(4 - 2\alpha_1 - \frac{\alpha_2}{\alpha_1^2} \right) m_2 m_1^2 \\ & - 2\left(4\alpha_1 + \pi - 4 - \frac{\alpha_2}{\alpha_1^2} \right) m_1^2 \tau^2 \Bigg\}.\end{aligned}$$

Finally,

$$\hat{m}_{f_2} = \frac{1}{n} S_n^{f_2} \to m_{f_2} := \frac{1}{\alpha_1}\left[(\alpha_1 - 2 + \pi)\mu_2 + (2-\pi)\mu_1^2\right]$$
$$= \frac{1}{\alpha_1}\left[(\alpha_1 - 2 + \pi)\left(m_2 - \tau^2\right) + (2-\pi)m_1^2\right] \text{ a.s.}$$

and

$$\sqrt{n}\left(\frac{1}{n} S_n^{f_2} - m_{f_2}\right) \to N\left(0, \gamma_{f_2}^2\right) \textit{ in distribution},$$

where

$$\begin{aligned}
\gamma_{f_2}^2 = \frac{1}{\alpha_1}\Bigg\{ & (\alpha_2 - 4\alpha_1 + 4 - \pi)m_4 \\
& + \left[2\alpha_2 - 11\alpha_1 - 2\left(\pi^3 - 5\pi^2 + 9\pi - 8\right) + (2-\pi)^2\frac{\alpha_2}{\alpha_1^2} + 2\pi(2-\pi)\frac{1}{\alpha_1}\right]\tau^4 \\
& + \left[4\alpha_1 - \alpha_2 + 2\pi^2 - 3\pi - 4 + 2\pi(2-\pi)\frac{1}{\alpha_1} + (2-\pi)^2\frac{\alpha_2}{\alpha_1^2}\right]m_2^2 \\
& + 2\left[10\alpha_1 - 2\alpha_2 + \pi^3 - 6\pi^2 + 12\pi - 12 - 2\pi(2-\pi)\frac{1}{\alpha_1} - (2-\pi)^2\frac{\alpha_2}{\alpha_1^2}\right]m_2\tau^2 \\
& + \left[2\left(\pi^2 + 4\pi - 8\right) + 2\pi(2-\pi)\frac{1}{\alpha_1} + (2-\pi)^2\frac{\alpha_2}{\alpha_1^2}\right]m_1^4 + 8[\alpha_1 + \pi - 2]m_1 m_3 \\
& - 2\left[4\alpha_1 + 2\left(\pi^2 + 3\pi - 8\right) + 2\pi(2-\pi)\frac{1}{\alpha_1} + (2-\pi)^2\frac{\alpha_2}{\alpha_1^2}\right]m_2 m_1^2 \\
& - 2\left[8\alpha_1 + \pi^3 - 6\pi^2 + 12\pi - 12 - 2\pi(2-\pi)\frac{1}{\alpha_1} - (2-\pi)^2\frac{\alpha_2}{\alpha_1^2}\right]m_1^2\tau^2\Bigg\}.
\end{aligned}$$

Notice that, by the definition in (7), the vectors $\{\mu_1, \ldots, \mu_{T-1}\}$ and $\{m_1, \ldots, m_{T-1}\}$ are linked by a linear transformation. We actually use the vectors $\mathbf{m} = \{1, m_1, \ldots, m_{T-1}\}$ and $\mu = \{1, \mu_1, \ldots, \mu_{T-1}\}$, which are related by

$$\mathbf{m} = U_{\tau^2} \cdot \boldsymbol{\mu}$$

with U_{τ^2} a $T \times T$ lower triangular matrix depending on τ^2 of the form

$$U_{\tau^2} = \begin{bmatrix} 1 & 0 & 0 & 0 & \ldots \\ 0 & 1 & 0 & 0 & \ldots \\ \tau^2 & 0 & 1 & 0 & \ldots \\ \ldots & & & & \end{bmatrix}.$$

More explicitly, one obtains from (7)

$$m_i = \sum_{\ell=0}^{i} \mathbf{I}_{(\ell = i \mod 2)} \rho_{\frac{i-\ell}{2}}(i) \tau^{i-\ell} \mu_\ell = \sum_{\ell=0}^{i} U_{\tau^2}[i, \ell]\mu_\ell \tag{8}$$

for $i = 1, \ldots, T-1$, where $\mathbf{I}_A$ indicates the indicator function of A. The matrix U_{τ^2} can be inverted, and the explicit inverse relations we will use in the following asymptotic theory are

$$\begin{aligned}
\boldsymbol{\mu}_1 &= m_1\\
\boldsymbol{\mu}_2 &= m_2 - \tau^2\\
\boldsymbol{\mu}_3 &= m_3 - 3m_1\tau^2\\
\boldsymbol{\mu}_4 &= m_4 - 6m_2\tau^2 + 3\tau^4.
\end{aligned}$$

Proof. The a.s. convergence of $\hat{m}_i$, for $i=1,\ldots,T-1$, $\hat{m}_{f1}$, and $\hat{m}_{f_2}$ follows from Proposition 1. It is only needed to compute an explicit expression for θ_{x^i}, which is easily obtained from Theorem 1 with $k=1$: for all $i=1,\ldots,T-1$

$$\begin{aligned}
\theta_{x^i} &= E\left[R_n^{x^i}\right]/E[N_n] = E[\beta_i(\lambda)]\\
&= E\left(\sum_{j=0}^{\lfloor i/2\rfloor}\rho_j(i)\lambda^{i-2j}\tau^{2j}\right)\\
&= \left(\sum_{j=0}^{i/2}\rho_j(i)\mu_{i-2j}\tau^{2j}\right) = m_i.
\end{aligned}$$

The corresponding expressions for θ_{f_1} and θ_{f_2} can be computed from the formulae for generic moments of functions $f: X^2 \to \mathbb{R}$ and $f: X^3 \to \mathbb{R}$ (see Appendix A), respectively. Using the relationship between $\boldsymbol{\mu}$ and $\mathbf{m}$, we obtain

$$\begin{aligned}
\theta_{f_1} = E\left[R_n^{f_1}\right]/E[N_n] &= \frac{1}{\alpha_1}\left[(\alpha_1-1)E\left[\beta_1^2(\lambda)\right] + (E[\beta_1(\lambda)])^2\right]\\
&= \frac{1}{\alpha_1}\left[(\alpha_1-1)E\left[\lambda^2\right] + (E[\lambda])^2\right]\\
&= \frac{1}{\alpha_1}\left[(\alpha_1-1)\boldsymbol{\mu}_2 + \boldsymbol{\mu}_1^2\right]\\
&= \frac{1}{\alpha_1}\left[(\alpha_1-1)\left(m_2-\tau^2\right) + m_1^2\right]
\end{aligned}$$

$$\begin{aligned}
\theta_{f_2} &= E\left[R_n^{f_2}\right]/E[N_n]\\
&= \frac{1}{\alpha_1}\left[(\alpha_1-2-\pi)E\left[\beta_1^2(\lambda)\right] + (1-\pi)(E[\beta_1(\lambda)])^2 + \pi(E[\beta_1(\lambda)])^2 + (1-\pi)(E[\beta_1(\lambda)])^2\right]\\
&= \frac{1}{\alpha_1}\left[(\alpha_1-2+\pi)E\left[\lambda^2\right] + (2-\pi)(E[\lambda])^2\right]\\
&= \frac{1}{\alpha_1}\left[(\alpha_1-2+\pi)\boldsymbol{\mu}_2 + (2-\pi)\boldsymbol{\mu}_1^2\right]\\
&= \frac{1}{\alpha_1}\left[(\alpha_1-2+\pi)\left(m_2-\tau^2\right) + (2-\pi)m_1^2\right].
\end{aligned}$$

The convergence in distribution of $\hat{m}_i$, for $i=1,\ldots,T-1$, $\hat{m}_{f_1}$, and $\hat{m}_{f_2}$ follows from Proposition 2, which also tells us how to calculate the variance of normal asymptotic distribution.

The asymptotic variances $\hat{m}_i$'s can be computed from Theorem 1:

$$\begin{aligned}
\gamma_i^2 = \chi_0 &= \frac{V\left[R_n^{x^i} - \theta_{x^i} N_n\right]}{E[N_n]} \\
&= \frac{1}{\alpha_1}\left[E\left[\left(R_n^{x^i} - \theta_{x^i} N_n\right)^2\right] - \left(E\left[R_n^{x^i} - \theta_{x^i} N_n\right]\right)^2\right] \\
&= \frac{1}{\alpha_1}\left[E\left[\left(R_n^{x^i} - \theta_{x^i} N_n\right)^2\right]\right] \\
&= \frac{1}{\alpha_1}\left[E\left[\left(R_n^{x^i}\right)^2\right] + \theta_{x^i}^2 E\left[N_n^2\right] - 2\theta_{x^i} E\left[N_n R_n^{x^i}\right]\right] \\
&= \frac{1}{\alpha_1}\left[E\left[\left(R_n^{x^i}\right)^2\right] + \alpha_2 m_i^2 - 2 m_i \alpha_2 E[\beta_i(\lambda)]\right] \\
&= \frac{1}{\alpha_1}\left[E\left[\left(R_n^{x^i}\right)^2\right] - \alpha_2 m_i^2\right] \\
&= \frac{1}{\alpha_1}\left[(\alpha_2 - \alpha_1) E\left[\beta_i^2(\lambda)\right] + \alpha_1 E[\beta_{2i}(\lambda)] - \alpha_2 m_i^2\right] \\
&= \frac{1}{\alpha_1}\left[(\alpha_2 - \alpha_1) E\left[\beta_i^2(\lambda)\right] + \alpha_1 m_{2i} - \alpha_2 m_i^2\right].
\end{aligned}$$

The asymptotic variances $\gamma_{f_1}^2$ and $\gamma_{f_2}^2$ require long calculations, which are sketched in Appendix A. □

By the definition of the μ_i's, $\boldsymbol{\mu} = V \cdot \mathbf{p}$ with V the Vandermonde matrix

$$V = \begin{bmatrix} 1 & 1 & 1 & \dots \\ b_1 & b_2 & b_3 & \dots \\ b_1^2 & b_2^2 & b_3^2 & \dots \\ \dots & & & \end{bmatrix} \tag{9}$$

of size $T \times T$. For vectors **a** and **b**, let (**a**,**b**) represent the concatenated vector. Then we can invert first moments and 2-autocovariances and 3-autocovariances as functions of the model parameters:

Lemma 2. *The* $\mathbb{R}^{T+2} \to \mathbb{R}^{T+2}$ *map*

$$\left(\mathbf{p}, \pi, \tau^2\right) \to \left(\mathbf{m},\, m_{f_1}, m_{f_2}\right),$$

is invertible, and its inverse is given by the continuous functions:

$$\pi = 1 - \frac{m_{f_2} - m_1^2}{m_{f_1} - m_1^2} \tag{10}$$

$$\tau^2 = m_2 - m_1^2 - \frac{\left(m_{f_1} - m_1^2\right)^2}{m_{f_2} - m_1^2} \tag{11}$$

$$\mathbf{p} = \mathbf{V}^{-1}\mathbf{U}^{-1}_{\tau^2\left(\mathbf{m},\mathbf{m}_{f_1},\mathbf{m}_{f_2}\right)}\cdot\mathbf{m}. \tag{12}$$

Proof The first two equalities are obtained by solving the system

$$\begin{cases} m_{f_1} = \dfrac{1}{\alpha_1}\left[(\alpha_1 - 1)\boldsymbol{\mu}_2 + \boldsymbol{\mu}_1^2\right] \\ m_{f_2} = \dfrac{1}{\alpha_1}\left[(\alpha_1 - 2 + \pi)\boldsymbol{\mu}_2 + (2-\pi)\boldsymbol{\mu}_1^2\right] \end{cases}$$

or, expressing the moments μ_i through sample moments m_i,

$$\begin{cases} m_{f_1} = \dfrac{1}{\alpha_1}\left[(\alpha_1 - 1)\left(m_2 - \tau^2\right) + m_1^2\right] \\ m_{f_2} = \dfrac{1}{\alpha_1}\left[(\alpha_1 - 2 + \pi)\left(m_2 - \tau^2\right) + (2-\pi)m_1^2\right] \end{cases}$$

with respect to the variable π and τ^2. Such inverses are continuous in the parameter range of the model because $m_{f_i} - m_1^2 = (1-\pi)^i\left(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1^2\right) \neq 0$ for $i = 1, 2$ being $\pi \neq 1$ by hypothesis and the states variance different from zero by the model (otherwise, we would have only one level).

The vector $\mathbf{p}$ is obtained inverting the system $\mathbf{m} = \mathbf{U}_{\tau^2}\boldsymbol{\mu} = \mathbf{U}_{\tau^2}\mathbf{V}\mathbf{p}$. Such inverse exists because U is a lower triangular matrix with all ones on the diagonal and V is a Vandermonde matrix with elements $b_i \neq b_j$ if $i \neq j$, for $i, j = 1, \ldots, T-1$, and it is continuous in the model parameters. □

The next theorem is the main result of our paper and gives the moment estimation of the model parameters:

Theorem 2. *If* $\{Y_\alpha\} = \left\{\left\{X_j^{(\Lambda_n)}\right\}_{j=1}^{N_n}\right\}$ *is a GFN-SLM, then*

$$\hat{\pi}_n = 1 - \frac{\hat{m}_{f_2} - \hat{m}_1^2}{\hat{m}_{f_1} - \hat{m}_1^2} \to \pi \text{ a.s.} \tag{13}$$

and

$$\sqrt{n}(\hat{\pi}_n - \pi) \to N\left(0, \gamma_\pi^2\right) \text{ in distribution},$$

where

$$\gamma_\pi^2 = J_F C_{(\hat{m}_1,\ldots,\hat{m}_{T-1})} J_F^t(1,1)$$

with

$$\mathbf{J_F} = \begin{bmatrix} \dfrac{2m_1\left(m_{f_1} - m_{f_2}\right)}{\left(m_{f_1} - m_1^2\right)^2} & 0 & \dfrac{m_{f_2} - m_1^2}{\left(m_{f_1} - m_1^2\right)^2} & -\dfrac{1}{m_{f_1} - m_1^2} \\ -\dfrac{2m_1\left(m_{f_1} - m_{f_2}\right)^2}{\left(m_{f_2} - m_1^2\right)^2} & 1 & -\dfrac{2\left(m_{f_1} - m_1^2\right)}{m_{f_2} - m_1^2} & \dfrac{\left(m_{f_1} - m_1^2\right)^2}{\left(m_{f_2} - m_1^2\right)^2} \end{bmatrix}.$$

and $C_{(\hat{m}_1,\ldots,\hat{m}_{T-1})}$ *explicitly calculable*:

$$C_{(\hat{m}_1,\ldots,\hat{m}_{T-1})}(t,r) = \frac{1}{\alpha_1}\{(\alpha_2-\alpha_1)E[\beta_t(\lambda)\beta_r(\lambda)] + \alpha_1 m_{t+r} - \alpha_2 m_t m_r\};$$
$$\hat{\tau^2}_n = \hat{m}_2 - \hat{m}_1^2 - \frac{\left(\hat{m}_{f_1} - \hat{m}_1^2\right)^2}{\hat{m}_{f_2} - \hat{m}_1^2} \to \tau^2 \text{ a.s.} \tag{14}$$

and

$$\sqrt{n}\left(\hat{\tau^2}_n - \tau^2\right) \to N\left(0, \gamma^2_{\tau^2}\right) \textit{ in distribution},$$

where

$$\gamma^2_{\tau^2} = J_F C_{(\hat{m}_1,\ldots,\hat{m}_{T-1})} J_F^t(2,2)$$

with $\mathbf{J}_F$ *and* $C_{(\hat{m}_1,\ldots,\hat{m}_{T-1})}$ *as above; finally*,

$$\hat{\mathbf{p}}_n = V^{-1} {U_{\hat{\tau^2}}}^{-1} \cdot \hat{\mathbf{m}}_n \to \mathbf{p} \text{ a.s.} \tag{15}$$

where $\hat{\mathbf{m}}_n = \{\mathbf{1}, \hat{m}_1, \ldots, \hat{m}_{T-1}\}$, and for $t = 1, \ldots, T$

$$\sqrt{n}(\hat{p}_t - p_t) \to N\left(0, \gamma^2_{p_t}\right) \textit{ in distribution},$$

where

$$\gamma^2_{p_t} = \mathbf{J}_G \mathbf{C}_{(\hat{m}_1,\ldots,\hat{m}_{T-1})} \mathbf{J}^{\mathbf{t}}_{\mathbf{G}}(\mathbf{t},\mathbf{t})$$

with $\boldsymbol{C}_{(\hat{m}_1\ldots,\hat{m}_{T-1})}$ *as above and* $\mathbf{J}_G$ *explicitly calculable in terms of the moments of the normal distribution, as indicated in the proof.*

Proof The a.e. convergences are simply a consequence of the a.e. convergence of $\left(\hat{\mathbf{m}}_n, \hat{m}_{f_1}, \hat{m}_{f_2}\right)$ to $\left(\mathbf{m},\ m_{f_1}, m_{f_2}\right)$ and the continuity of the functions in the previous Lemma.

By Lemma 1, we know the asymptotics of the statistics $\hat{\mathbf{m}}_n$, $\hat{m}_{f_1}$, and $\hat{m}_{f_2}$ so that the asymptotic variances of the present theorem follow from a multidimensional delta method as follows. We evaluate all functions in the asymptotic values $\mathbf{m}_n$, m_{f_1}, and m_{f_2} of $\hat{\mathbf{m}}_n$, $\hat{m}_{f_1}$, and $\hat{m}_{f_2}$.

In order to derive the asymptotic variances of π and τ^2, we consider the function

$$F : \mathbb{R}^4 \to \mathbb{R}^2$$
$$\left(\hat{m}_1, \hat{m}_2, \hat{m}_{f_1}, \hat{m}_{f_2}\right) \mapsto \left(\hat{\pi}_n = 1 - \frac{\hat{m}_{f_2} - \hat{m}_1^2}{\hat{m}_{f_1} - \hat{m}_1^2}, \hat{\pi^2}_n = \hat{m}_2 - \hat{m}_1^2 - \frac{\left(\hat{m}_{f_1} - \hat{m}_1^2\right)^2}{\hat{m}_{f_2} - \hat{m}_1^2}\right)$$

whose Jacobian calculated in the asymptotic value of the vector $\left(\widehat{m}_1, \widehat{m}_2, \widehat{m}_{f_1}, \widehat{m}_{f_2}\right)$ is given by

$$\mathbf{J_F} = \begin{bmatrix} \dfrac{2m_1\left(m_{f_1} - m_{f_2}\right)}{\left(m_{f_1} - m_1^2\right)^2} & 0 & \dfrac{m_{f_2} - m_1^2}{\left(m_{f_1} - m_1^2\right)^2} & -\dfrac{1}{m_{f_1} - m_1^2} \\ -\dfrac{2m_1\left(m_{f_1} - m_{f_2}\right)^2}{\left(m_{f_2} - m_1^2\right)^2} & 1 & -\dfrac{2\left(m_{f_1} - m_1^2\right)}{m_{f_2} - m_1^2} & \dfrac{\left(m_{f_1} - m_1^2\right)^2}{\left(m_{f_2} - m_1^2\right)^2} \end{bmatrix}.$$

If we denote with $C_{(\hat{m}_1,\hat{m}_2,\hat{m}_{f_1},\hat{m}_{f_2})}$ the covariance matrix of the vector $(\hat{m}_1,\hat{m}_2,\hat{m}_{f_1},\hat{m}_{f_2})$, we have

$$\begin{aligned}
\boldsymbol{C}_{(\hat{m}_1,\hat{m}_2,\hat{m}_{f_1},\hat{m}_{f_2})}(1,1) &= \gamma_1^2\\
\boldsymbol{C}_{(\hat{m}_1,\hat{m}_2,\hat{m}_{f_1},\hat{m}_{f_2})}(2,2) &= \gamma_2^2\\
\boldsymbol{C}_{(\hat{m}_1,\hat{m}_2,\hat{m}_{f_1},\hat{m}_{f_2})}(3,3) &= \gamma_{f_1}^2\\
\boldsymbol{C}_{(\hat{m}_1,\hat{m}_2,\hat{m}_{f_1},\hat{m}_{f_2})}(4,4) &= \gamma_{f_2}^2
\end{aligned}$$

with the expression of the variances given in Lemma 1.

The off-diagonal terms are explicitly computed in Appendix A.

Using the multidimensional delta method, the variances of $\hat{\pi}$ and $\hat{\tau^2}$ are the diagonal terms of the matrix $\mathbf{J_F}\mathbf{C}_{(m_1 m_2 m_{f_1} m_{f_2})}\mathbf{J_F^t}$, that is

$$\begin{aligned}
\gamma_\pi^2 = {} & \frac{4m_1^2\left(m_{f_1}-m_{f_2}\right)^2}{\left(m_{f_1}-m_1^2\right)^4}\gamma_1^2 + \frac{4m_1\left(m_{f_1}-m_{f_2}\right)\left(m_{f_2}-m_1^2\right)}{\left(m_{f_1}-m_1^2\right)^4}\mathrm{Cov}\left[\hat{m}_1,\hat{m}_{f_1}\right]\\
& - \frac{4m_1\left(m_{f_1}-m_{f_2}\right)}{\left(m_{f_1}-m_1^2\right)^3}\mathrm{Cov}\left[\hat{m}_1,\hat{m}_{f_2}\right] + \frac{\left(m_{f_2}-m_1^2\right)^2}{\left(m_{f_1}-m_1^2\right)^4}\gamma_{f_1}^2\\
& - \frac{2\left(m_{f_2}-m_1^2\right)}{\left(m_{f_1}-m_1^2\right)^3}\mathrm{Cov}\left[\hat{m}_{f_1},\hat{m}_{f_2}\right] + \frac{1}{\left(m_{f_1}-m_1^2\right)^2}\gamma_{f_2}^2\\
\gamma_{\tau^2}^2 = {} & \frac{4m_1^2\left(m_{f_1}-m_{f_2}\right)^4}{\left(m_{f_2}-m_1^2\right)^4}\gamma_1^2 - \frac{4m_1\left(m_{f_1}-m_{f_2}\right)^2}{\left(m_{f_2}-m_1^2\right)^2}\mathrm{Cov}[\hat{m}_1,\hat{m}_2]\\
& + \frac{8m_1\left(m_{f_1}-m_{f_2}\right)^2\left(m_{f_1}-m_1^2\right)}{\left(m_{f_2}-m_1^2\right)^3}\mathrm{Cov}\left[\hat{m}_1,\hat{m}_{f_1}\right]\\
& - \frac{4m_1\left(m_{f_1}-m_{f_2}\right)^2\left(m_{f_1}-m_1^2\right)^2}{\left(m_{f_2}-m_1^2\right)^4}\mathrm{Cov}\left[\hat{m}_1,\hat{m}_{f_2}\right]\\
& + \gamma_2^2 - \frac{4\left(m_{f_1}-m_1^2\right)}{\left(m_{f_2}-m_1^2\right)}\mathrm{Cov}\left[\hat{m}_2,\hat{m}_{f_1}\right] + \frac{2\left(m_{f_1}-m_1^2\right)^2}{\left(m_{f_2}-m_1^2\right)^2}\mathrm{Cov}\left[\hat{m}_2,\hat{m}_{f_2}\right]\\
& + \frac{4\left(m_{f_1}-m_1^2\right)^2}{\left(m_{f_2}-m_1^2\right)^2}\gamma_{f_1}^2 - \frac{4\left(m_{f_1}-m_1^2\right)^3}{\left(m_{f_2}-m_1^2\right)^3}\mathrm{Cov}\left[\hat{m}_{f_1},\hat{m}_{f_2}\right] + \frac{4\left(m_{f_1}-m_1^2\right)^4}{\left(m_{f_2}-m_1^2\right)^4}\gamma_{f_2}^2.
\end{aligned}$$

For the variances vector $\gamma_{\mathbf{p}}^2$, we have to consider the function

$$\begin{array}{rccc}
G: & \mathbb{R}^{T-1} & \to & \mathbb{R}^T \\
 & (\hat{m}_1,\ldots,\hat{m}_{T-1}) & \mapsto & \hat{\mathbf{p}}
\end{array} = V^{-1}U_{\hat{\tau^2}}^{-1}\hat{\mathbf{m}}.$$

If we denote with $\mathbf{J_G}$ the Jacobian matrix of the function G evaluated in $(m_1,\ldots,m_{T-1})$ and with $\mathrm{C}_{(\hat{m}_1,\ldots,\hat{m}_{T-1})}$ the variance–covariance matrix of the vector $(\hat{m}_1,\ldots,\hat{m}_{T-1})$,

then the main diagonal of the matrix $\mathbf{J_G C}_{(\hat{m}_1,\ldots,\hat{m}_{T-1})}\mathbf{J_G^t}$ consists of the variances $\gamma^2_{p_t}$, for $t = 1, \ldots, T-1$.

Notice that for $t = 1, \ldots, T-1$ we have $\mathbf{C}_{(\hat{m}_1,\ldots,\hat{m}_{T-1})}(t,t) = \gamma_t^2$, whose expression is given in Lemma 1, whereas for $t, r = 1, \ldots, T-1$ with $t \neq r$, we have

$$\begin{aligned} C_{(\hat{m}_1\ldots,\hat{m}_{T-1})}(t,r) &= \mathrm{Cov}[\hat{m}_t, \hat{m}_r] \\ &= \frac{1}{\alpha_1}\{(\alpha_2 - \alpha_1)E[\beta_t(\lambda)\beta_r(\lambda)] + \alpha_1 m_{t+r} - \alpha_2 m_t m_r\}. \end{aligned}$$

□

We end this section by observing that the variances given by the previous theorem allow us to obtain CIs for segmentation parameters. Denoting by α the confidence level and considering the normal asymptotic distribution of estimators, we can derive the following CIs for π, τ^2, and $\mathbf{p}$, respectively:

$$\left(\hat{\pi}_n - z_{\alpha/2}\sqrt{\frac{\gamma^2_\pi}{n}}, \hat{\pi}_n + z_{\alpha/2}\sqrt{\frac{\gamma^2_\pi}{n}}\right)$$

$$\left(\hat{\tau^2_n} - z_{\alpha/2}\sqrt{\frac{\gamma^2_{\tau^2}}{n}}, \hat{\tau^2_n} + z_{\alpha/2}\sqrt{\frac{\gamma^2_{\tau^2}}{n}}\right)$$

$$\left(\hat{p_i} - z_{\alpha/2}\sqrt{\frac{\gamma^2_{p_i}}{n}}, \hat{p_i} + z_{\alpha/2}\sqrt{\frac{\gamma^2_{p_i}}{n}}\right) \; i = 1, \ldots, T.$$

4 A discrete model for array-CGH data

Array-CGH is a microarray technology that allows one to detect and map genomic alterations. The goal of array-CGH analysis is to identify the boundaries of the regions where the number of DNA copies changes and then to label each region as loss, neutral, or gain. The genomic profile obtained from an array-CGH experiment can be considered as a signal made of noisy segments with different lengths and with mean levels that shift their values according to the DNA copy number.

In the mathematical model of MAGI *et al.* (2010), this signal has been considered as generated by the sum of two processes: a biological process due to a real variation of the number of DNA copies and a white noise process that mimics experimental error. We thus consider sequential observations $\mathbf{Y} = (Y_1, \ldots, Y_N)$ to be realizations of the sum of two independent stochastic processes:

$$Y_i = \Lambda_i + \varepsilon_i$$

where ε_i is normally distributed white noise with variance τ^2, $\varepsilon_i \sim N(0, \sigma^2_\varepsilon)$, and Λ_i is the unobserved mean level.

In MAGI *et al.* (2010), the Λ_i's have been taken to be normally distributed with the values taken in a specific sample estimated during the statistical analysis; however, we make here the additional observation that these values are not arbitrary, as they reflect the possible values of (the log of) the fluorescent ratio of DNA copy number against absence of aberration. For deleted regions, the normalized $\log_2$-ratio is $\log_2(1/2) = -1$, whereas for amplified regions, the normalized $\log_2$-ratio is $\log_2(3/2) = 0.5849$ or $\log_2(4/2) = 1$ for four copies amplification. The value 0 corresponds to no aberrations. Hence, the possible states of Λ_i can be determined at the start of the analysis and chosen to be taken from a finite distribution on $b = \{b_1, \ldots, b_T\}$ with parameters $\mathbf{p} = \{p_1, \ldots, p_T\}$. To avoid missing unusual values, one can, as we actually do, insert additional values of the b_j's with probability 0: this will be recognized during the analysis and thus such states can be later removed. We then believe that, as long as the relevant biologically justified values are considered, simple variations in the choice of the vector b are not likely to alter the statistical analysis we are going to perform; we verified such claim with a systematic investigation of the synthetic Lai *et al.* data set by using different choices of the state vector b (Section 6).

Then we consider the process $\left\{X_i^{(\lambda)}\right\}$, whose elements are given by

$$X_i^{(\lambda)} = \lambda + \varepsilon_i,$$

which corresponds to the process $\{Y_i\}$, with the fixed value $\Lambda_i = \lambda$. The random variables $X_i^{(\lambda)}$ are i.i.d., and as the stochastic processes Λ_i and ε_i are independent, we obtain that $E[Y_i] = E[\Lambda_i] = \mu_1$, $\mathrm{Var}[Y_i] = V[\Lambda_i] + \tau^2$, and consequently

$$\begin{aligned} E\left[X_i^{(\lambda)}\right] &= \lambda \\ V\left[X_i^{(\lambda)}\right] &= \tau^2; \end{aligned}$$

therefore,

$$X_i^{(\lambda)} \sim N\left(\lambda, \tau^2\right).$$

Sequences of observations of given lengths with the same mean correspond to chromosomal aberrations, and their lengths N_i's have been taken in MAGI *et al.* (2010) to be i.i.d. geometrically distributed stochastic process $N_i \sim \mathcal{G}(\pi)$, with mean π, independent from the Λ_j's and ε_j's. As pointed out by an anonymous referee, this might not be a very appropriate model in a number of cases in which high amplitude gains are often of small genomic size: in such a case, the N_i's would no longer be identically distributed and the parameter π should depend on Λ_i. However, this is not case in various other situations, including primarily cancer genomic analysis (Bayani et al. (2007)). In addition to this, the assumption of constant π simplifies the mathematical analysis while producing very good results in terms of segmentation (Section 6). For these reasons, we stick to the assumption that the N_i's are i.i.d. with $N_i \sim \mathcal{G}(\pi)$. A more general

method of moments than the one described here could very likely be able to deal with varying π's, but this requires extensions of the mathematical results, and we are currently investigating such possibility.

With the assumptions made so far, the data originated by an array-CGH experiment can be described through the GFN-SLP with

$$\{N_i\} \sim G(\pi) \text{ i.i.d.}$$

$$\{\Lambda_i\} \sim F(p_1, \ldots, p_T) \text{ i.i.d.}$$

$$\left\{X_i^{(\lambda)}\right\} \sim \mathrm{N}(\lambda, \tau^2) \text{ i.i.d.}$$

where the processes are mutually independent.

5 GFN-SLP analysis and segmentation of array-CGH data

From the results of the previous sections, we have an algorithm to estimate the parameter vector $\{\pi,\mathbf{p},\tau^2\} = \{\pi,p_1,\ldots,p_T,\tau^2\}$ of the aforementioned model once assigned the state vector $\{b_1,\ldots,b_T\}$. The main difference with existing estimation methods is that we can estimate all the parameters $\{\pi,p_1,\ldots,p_T,\tau^2\}$ in one step, whereas most methods require assigning some of the parameters and often need iterative steps. For this reason, our method is likely to be faster than any other currently available algorithm (Section 7).

Collecting formulae for reader's convenience, the method consists of evaluating

$$\begin{aligned}
\hat{m}_i &= \frac{1}{n} S_n^{x^i}, \qquad i = 1, \ldots, T-1 \\
\hat{m}_{f_1} &= \frac{1}{n} S_n^{x_1 \cdot x_2} \\
\hat{m}_{f_2} &= \frac{1}{n} S_n^{x_1 \cdot x_3}.
\end{aligned}$$

as in (6) and $\hat{\mathbf{m}}_n = \{1, \hat{m}_1, \ldots, \hat{m}_{T-1}\}$ from the data. Then the GFN-SLP parameter estimators based on the method of moments are

$$\hat{\pi} = 1 - \frac{\hat{m}_{f_2} - \hat{m}_1^2}{\hat{m}_{f_1} - \hat{m}_1^2}$$

from (13),

$$\hat{\tau^2} = \hat{m}_2 - \hat{m}_1^2 - \frac{\left(\hat{m}_{f_1} - \hat{m}_1^2\right)^2}{\hat{m}_{f_2} - \hat{m}_1^2}$$

from (14) and

$$\hat{\mathbf{p}} = V^{-1} U_{\hat{\tau^2}}^{-1} \cdot \hat{\mathbf{m}}_n$$

from (15) with $U_{\hat{\tau^2}}$ from (8) and V from (9).

Once the parameter estimation is performed, the segmentation can be completed by some of the existing methods. In the following simulations, we apply once a Viterbi algorithm based on the HMM representation of the GFN-SLM. Following SALAS AND BOES (1980),

$$\Lambda_i = (1 - z_{i-1})\Lambda_{i-1} + z_{i-1}(\boldsymbol{\mu}_1 + \delta_i),$$

where

- $z_1, z_2, \ldots$ are i.i.d. random variables taking the values 0, 1 with probabilities $P_\pi[z_i = 1] = \pi$ and $P_\pi[z_i = 0] = 1 - \pi$.
- $\delta_1, \delta_2, \ldots$ are i.i.d. random variables with finite distribution $F(p_1, \ldots, p_T)$,

which is a one-step Markov chain with initial distribution $\mathbf{p} = \{p_1, \ldots, p_T\}$, and transition matrix is $\mathbf{P} = \left\{\mathbf{P}_{ij}\right\}_{\mathbf{i,j}=1}^{\mathbf{T}}$ given by

$$P_{ij} = P\left[\Lambda_t = b_j | \Lambda_{t-1} = b_i\right] = \begin{cases} (1 - \pi) + \pi p_j & i = j \\ \pi p_j & i \neq j \end{cases} \quad (16)$$

and emission matrix is $\mathbf{E} = \left\{\mathbf{E}_{\mathbf{b_k y_j}}\right\}$, with

$$E_{b_k y_j} = P\left[Y_t = y_j | \Lambda_t = \lambda_k\right] = \frac{e^{-\frac{\left(y_j - b_k\right)^2}{2\tau^2}}}{\sqrt{2\pi\tau^2}}. \quad (17)$$

Some tests have been performed, and results are presented below. All figures show the segmentations (black lines) over the observed $\log_2$-ratio (light gray point). X axis runs along the entire genome, according to the physical mapping.

The first test has been performed on the data set V22711-4Q provided by the Diagnostic Genetic Unit, Careggi Hospital, University of Firenze, consisting of approximately 44 000 clones and a very noisy signal. To mitigate noise, we used the waves aCGH correction or WACA algorithm (Lepretre et al., 2010) to de-wave the signal.

The state vector we gave as input contains values that are equispaced and symmetric around the origin: $\{-2.1, -1.8, -1.5, -1.2, -0.9. -0.6, -0.3, 0, 0.3, 0.6, 0.9, 1.2, 1.5, 1.8, 2.1\}$. This certainly contains extra states, but it is likely to contain all states of interest. Parameters were subsequently estimated at

$$\begin{aligned} \hat{\pi} &= 0.02830132 \\ \hat{\tau}^2 &= 0.04843138 \\ \{\hat{p}_1, \ldots, \hat{p}_{15}\} &= \{0, 0, 0.0003295399, 0.0003663509, 0, 0.0111484618, \\ &\quad 0.3862741556, 0.2337560497, 0.3038405335, 0.0626967030, \\ &\quad 0, 0.0010005048, 0.0005817476, 0, 0\}, \end{aligned}$$

and the resulting segmentation is shown in Figure 1. Notice that many states have been indicated to have negligible probability. With a cutoff at 1%, only five states remain.

In particular, we can focus on the first chromosome, made of about 4000 clones, to see better what happens in detail (Figure 2).

The same data, analyzed with the SLM algorithm (see Magi *et al*. (2010)), freely available on R environment, produce the segmentation shown in Figure 3; and highlighting the first chromosome as before, we obtain the segmentation in Figure 4.

The second test has been performed on the genomic profile of chromosome 7 in sample GBM29 of the BREDEL *et al*. (2005) data set; the results are plotted in Figure 5 together with the SLM segmentation. Figure 5b shows that GFN-SLP is not able to correctly estimate the value of the state at the extremes. However, the principal aim of a segmentation method is to predict the breakpoints of each segment. In fact, the fine estimation of the level of each state may be assessed by the usage of array-CGH calling methods, such as FastCall (BENELLI *et al*., 2010) or CGHcall (VAN DE WIEL *et al*., 2007).

A comparison with other segmentations of the same data set appears in MAGI *et al*. (2010).

Numerical tests seem to indicate that our estimation method is quite sensitive, as it identifies even small CNV regions, which are overlooked by other methods. The main reason is the size of the estimated π, which is generally larger than other values usually adopted. Nonetheless, our method is able to identify large deletions or amplifications.

The asymptotic results of Theorem 2 allow, in principle, to write CIs for the parameters. This is a relevant difference with other estimation methods, but its application requires some accurate estimates on the sample size in order to be able

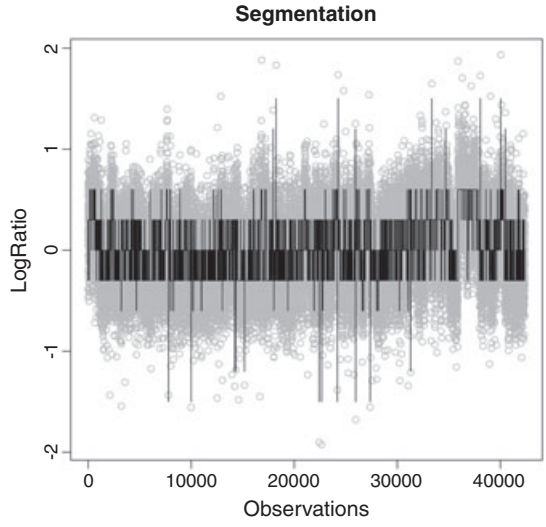


Fig. 1. GFN segmentation of V22711-4Q data along the entire genome.

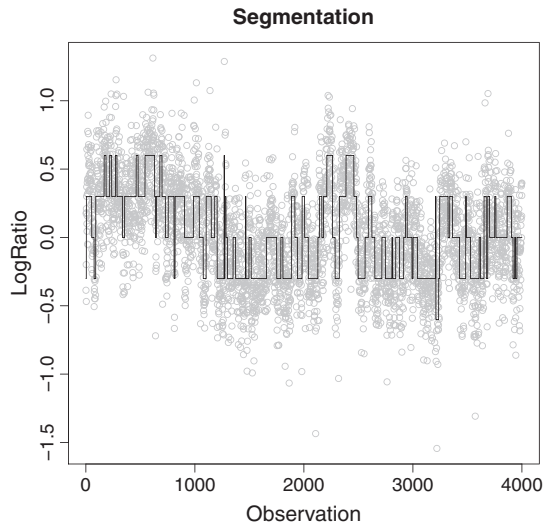


Fig. 2. GFN segmentation of the first chromosome of V22711-4Q data.

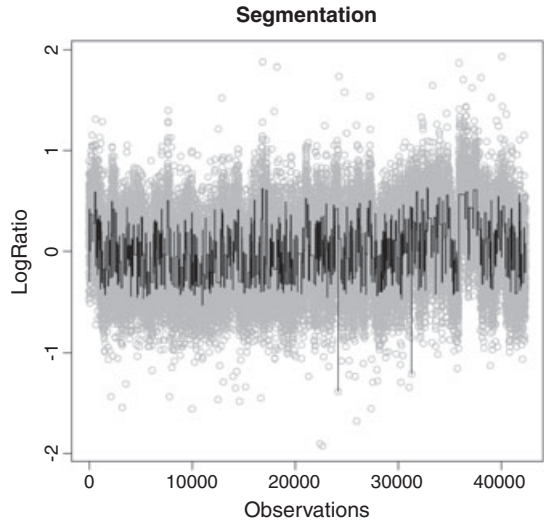


Fig. 3. SLM segmentation of V22711-4Q data along the entire genome.

to guarantee applicability of asymptotic variances. We do not pursue this direction in this paper, but simply show one numerical test on the data set V22711-4Q, whose point estimations are listed in the succeeding paragraph. Only CIs for π and τ are meaningful, as the CIs for the p_i's are too wide. Results are reported in Table 1.

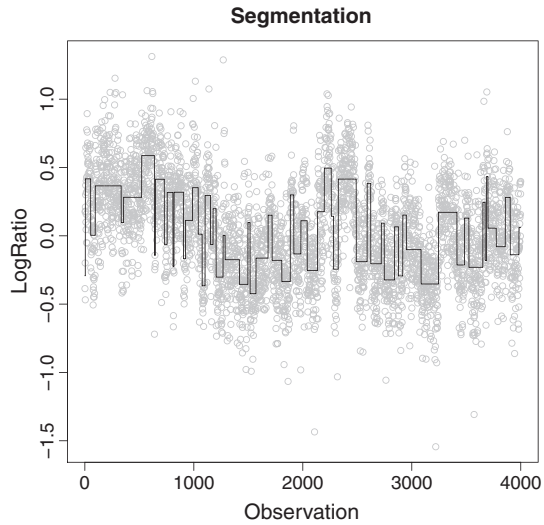
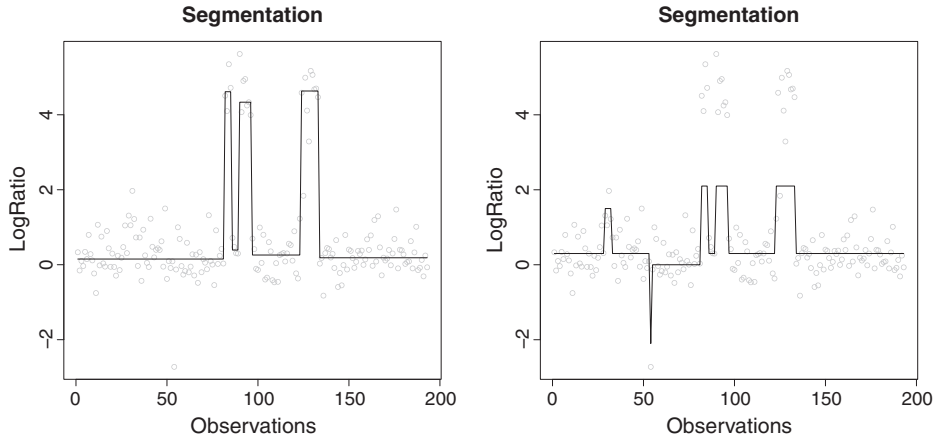


Fig. 4. SLM segmentation of the first chromosome of V22711-4Q data.



(a) SLM segmentation (b) GFN segmentation

Fig. 5. Comparison between the SLM and GFN segmentations on genomic profile of chromosome 7 in sample GBM29 of BREDEL *et al.* (2005) data set.

6 Comparison with state-of-the-art algorithms

To estimate the accuracy of the GFN-SLP algorithm in identifying the aberrations at the boundaries, we applied our algorithm on the synthetic chromosomes generated by LAI *et al.* (2005) (the data are freely available for download at http://www.chip.org/~ppark/Supplements/Bioinformatics05b.html).

The LAI *et al.* (2005) data set is made of synthetic chromosomes with four different aberration widths (5, 10, 20, and 40 probes) and four different signal-to-noise ratio (SNR) levels (1, 2, 3, and 4). For each aberration width and SNR, there are 100 independently simulated chromosomes with 100 probes in total. Here, we considered the most challenging situation where SNR = 1 and SNR = 2.

We applied GFN-SLP with four different settings of state vector, all with 15 equally spaced states taken between the indicated extremes:

- $b1 = \{-1.0, \ldots, 1.0\}$
- $b2 = \{-1.5, \ldots, 1.5\}$
- $b3 = \{-2.1, \ldots, 2.1\}$
- $b4 = \{-2.5, \ldots, 2.5\}$

We also used the circular binary segmentation (CBS; OLSHEN *et al.*, 2004), SLM (MAGI *et al.*, 2010), and HMM (Fridlyand et al., 2004) methods on these data, and we compared their performance by generating the ROC curve. To generate ROC curves, we calculated the true-positive rates (TPRs) and the false-positive rates (FPRs) as in LAI *et al.* (2005). TPR is defined as the number of probes inside the aberration whose fitted values are above the threshold level divided by the number of probes in the aberration. FPR is defined as the number of probes outside the aberration whose fitted values are above the threshold level divided by the total number of probes outside the aberration. The results reported in Figure 6 demonstrate that GFN-SLP outperforms the HMM algorithm in terms of both sensitivity and specificity and obtains comparable performance than the CBS method. The SLM algorithm by MAGI *et al.* (2010) outperforms the other methods for both SNR = 1 and SNR = 2. The simulation study we performed also indicates that changing the state vector b has little effect on the global performance of our GFN-SLP algorithm.

7 Computational performance

A common drawback of segmentation algorithms is the long running time required to segment real high-density arrays. The rapid growth of microarray size and resolution

Table 1. Confidence intervals for π and τ on data set V22711-4Q

Parameter	CI lower bound	CI upper bound
π	(99%) 0.0195	(99%) 0.0370
τ	(99%) 0.0472	(99%) 0.0496
π	(95%) 0.0216	(95%) 0.0349
τ	(95%) 0.0474	(95%) 0.0493

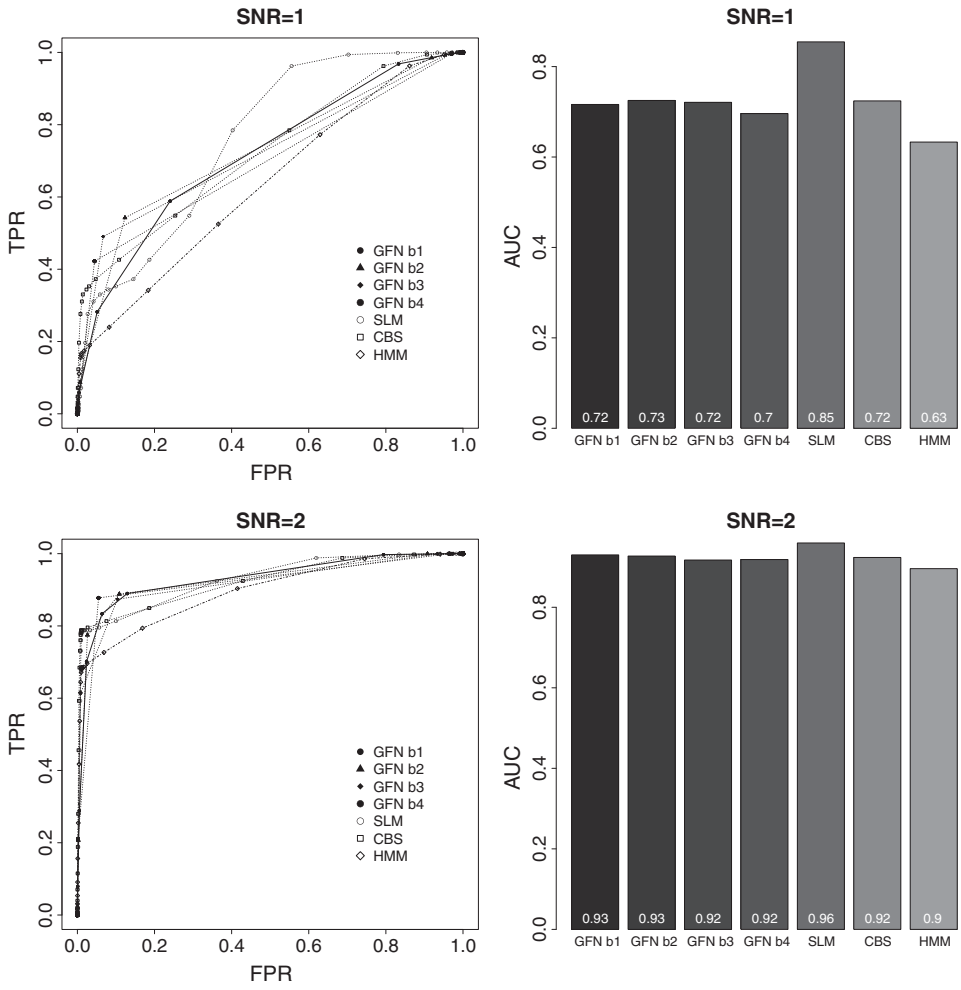


Fig. 6. ROC curves and area under the curve bar plot for GFN, CBS, SLM, and HMM on the synthetic chromosomes of Lai *et al*. (2005) data set.

requires segmentation algorithms with high computational performance. For this reason, we have tested the speed of GFN-SLP algorithm through an extensive experimentation on synthetic chromosomes and have compared its performance with respect to that of the other three methods. To this end, we generated synthetic chromosomes with different numbers of alterations (from 1 to 10) and different SNR (from 1 to 4).

We have tested the computational performances of the three algorithms on chromosomes with sizes from 500 to 32,000 clones (and with aberration width fixed to 30 clones).

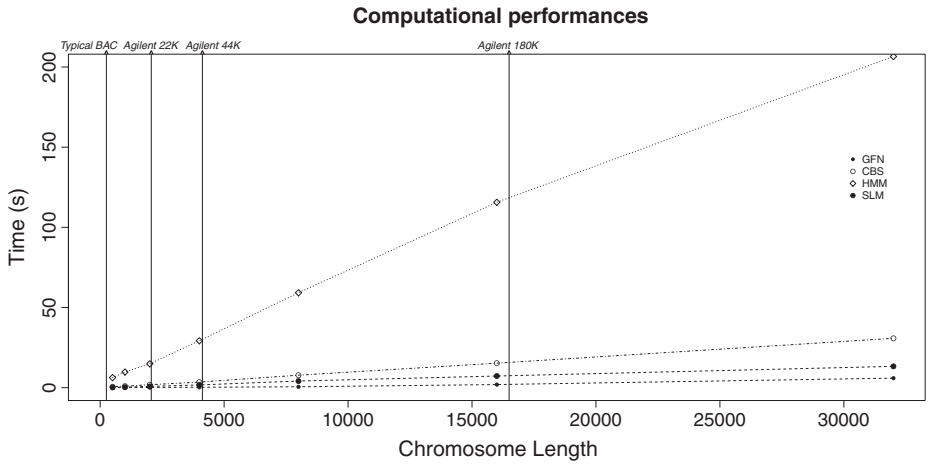


Fig. 7. Computational time comparison (in seconds) between GFN, CBS, SLM, and HMM methods. Each value of the table is calculated by averaging the times taken by each algorithm to segment synthetic chromosomes. We compared all the methods on chromosomes with sizes that ranges from 500 to 32,000 clones.

The results of all the simulations are summarized in Figure 7. Each value of the figure is calculated by averaging the times taken by each algorithm to segment the chromosomes: GFN-SLP outperforms the other three segmentation methods.

Appendix A. Calculating moments

Here is the calculation of generic moments required by Propositions 1 and 2, in terms of the moments of N_n and R_n^f. We provide formulae (without derivation) also reported in Cline's paper, for $f: X \to \mathbb{R}$ and $f: X^2 \to \mathbb{R}$, and add formulae for $f: X^3 \to \mathbb{R}$, which will be used in the proof of Lemma 1. As in Cline, we do this under the following assumption:

1. $\left\{X_j^{(\lambda)}\right\}$, $\lambda \in \mathbf{\Lambda}$, is a family of independent stochastic processes, each of which is a sequence of exchangeable random elements of $\mathbf{X}$.
2. $\{N_n\}$ and $\{\Lambda_n\}$ are sequences of i.i.d. random elements of $\mathbf{N}$ and Λ, respectively, and are independent of each other and of $\left\{X_j^{(\lambda)}\right\}$

Here are the formulae:
(1) for $f : \mathbf{X} \to \mathbb{R}$

$$E\left[R_n^f\right] = E[N_n]E\left[f\left(X_1^{(\Lambda_n)}\right)\right]$$

$$V\left[R_n^f - \theta N_n\right] = E[N_n]V\left[f\left(X_1^{(\Lambda_n)}\right)\right] + E[N_n(N_n-1)]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)}\right), f\left(X_2^{(\Lambda_n)}\right)\right]$$

We observe that in this case the autocovariances between $R_i^{f-\theta}$ are all zero.

(2) for $f : \mathbf{X}^2 \to \mathbb{R}$.

$$E\left[R_n^f\right] = E[N_n - 1]E\left[f\left(X_1^{(\Lambda_n)}, X_2^{(\Lambda_n)}\right)\right] + E\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}\right)\right].$$

$$\begin{aligned} V\left[R_n^f - \theta N_n\right] = {} & E[N_n - 1]V\left[f\left(X_1^{(\Lambda_n)}, X_2^{(\Lambda_n)}\right)\right] \\ & + 2E\left[(N_n - 2)_+\right]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)}, X_2^{(\Lambda_n)}\right), f\left(X_2^{(\Lambda_n)}, X_3^{(\Lambda_n)}\right)\right] \\ & + E\left[(N_n - 2)_+(N_n - 3)_+\right]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)}, X_2^{(\Lambda_n)}\right), f\left(X_3^{(\Lambda_n)}, X_4^{(\Lambda_n)}\right)\right] \\ & + V\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}\right)\right] \\ & + 2P[N_n > 1]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)}, X_2^{(\Lambda_n)}\right), f\left(X_2^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}\right)\right] \\ & + 2E\left[(N_n - 2)_+\right]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)}, X_2^{(\Lambda_n)}\right), f\left(X_3^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}\right)\right] \\ & + V[N_n]\left(E\left[f\left(X_1^{(\Lambda_n)}, X_2^{(\Lambda_n)}\right) - \theta\right]\right)^2. \end{aligned}$$

The autocovariances between $R_i^{f-\theta}$ are all zero, except when $R_i^{f-\theta}$ are consecutive elements of the process.

$$\begin{aligned} \mathrm{Cov}\left[R_n^{f-\theta}, R_{n+1}^{f-\theta}\right] = {} & P[N_{n+1} = 1]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}\right), f\left(X_1^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}\right)\right] \\ & + P[N_{n+1} > 1]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}\right), f\left(X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}\right)\right] \\ & + E\left[(N_{n+1} - 2)_+\right]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}\right), f\left(X_2^{(\Lambda_{n+1})}, X_3^{(\Lambda_{n+1})}\right)\right] \\ & + P[N_{n+1} > 1]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}\right), f\left(X_2^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}\right)\right] \end{aligned}$$

(3) for $f : \mathbf{X}^3 \to \mathbb{R}$

$$\begin{aligned} E\left[R_n^f\right] = {} & E\left[(N_n - 2)_+\right]E\left[f\left(X_1^{(\Lambda_n)}, X_2^{(\Lambda_n)}, X_3^{(\Lambda_n)}\right)\right] \\ & + P[N_n > 1]E\left[f\left(X_1^{(\Lambda_n)}, X_2^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}\right)\right] \\ & + P[N_{n+1} = 1]E\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}\right)\right] \\ & + P[N_{n+1} > 1]E\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}\right)\right] \end{aligned}$$

$$
\begin{aligned}
V\left[R_n^{f-\theta}\right] &= E\left[(N_n-2)_+\right]V\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_3^{(\Lambda_n)}\right)\right]\\
&+2E\left[(N_n-3)_+\right]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_3^{(\Lambda_n)}\right),f\left(X_2^{(\Lambda_n)},X_3^{(\Lambda_n)},X_4^{(\Lambda_n)}\right)\right]\\
&+2E\left[(N_n-4)_+\right]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_3^{(\Lambda_n)}\right),f\left(X_3^{(\Lambda_n)},X_4^{(\Lambda_n)},X_5^{(\Lambda_n)}\right)\right]\\
&+2E\left[(N_n-4)_+\right]E\left[(N_n-5)_+\right]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_3^{(\Lambda_n)}\right),f\left(X_4^{(\Lambda_n)},X_5^{(\Lambda_n)},X_6^{(\Lambda_n)}\right)\right]\\
&+P[N_n>1]V\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_1^{(\Lambda_{n+1})}\right)\right]\\
&+P[N_{n+1}=1]V\left[f\left(X_1^{(\Lambda_n)},X_1^{(\Lambda_{n+1})},X_1^{(\Lambda_{n+2})}\right)\right]\\
&+P[N_{n+1}>1]V\left[f\left(X_1^{(\Lambda_n)},X_1^{(\Lambda_{n+1})},X_2^{(\Lambda_{n+1})}\right)\right]\\
&+2P[N_n>2]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_3^{(\Lambda_n)}\right),f\left(X_2^{(\Lambda_n)},X_3^{(\Lambda_n)},X_1^{(\Lambda_{n+1})}\right)\right]\\
&+2P[N_n>3]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_3^{(\Lambda_n)}\right),f\left(X_3^{(\Lambda_n)},X_4^{(\Lambda_n)},X_1^{(\Lambda_{n+1})}\right)\right]\\
&+2E\left[(N_n-4)_+\right]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_3^{(\Lambda_n)}\right),f\left(X_4^{(\Lambda_n)},X_5^{(\Lambda_n)},X_1^{(\Lambda_{n+1})}\right)\right]\\
&+2P[N_{n+1}=1]P[N_n>2]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_3^{(\Lambda_n)}\right),f\left(X_3^{(\Lambda_n)},X_1^{(\Lambda_{n+1})},X_1^{(\Lambda_{n+2})}\right)\right]\\
&+2P[N_{n+1}=1]E\left[(N_n-3)_+\right]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_3^{(\Lambda_n)}\right),f\left(X_4^{(\Lambda_n)},X_1^{(\Lambda_{n+1})},X_1^{(\Lambda_{n+2})}\right)\right]\\
&+2P[N_{n+1}>1]P[N_n>2]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_3^{(\Lambda_n)}\right),f\left(X_3^{(\Lambda_n)},X_1^{(\Lambda_{n+1})},X_2^{(\Lambda_{n+1})}\right)\right]\\
&+2P[N_{n+1}>1]E\left[(N_n-3)_+\right]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_3^{(\Lambda_n)}\right),f\left(X_4^{(\Lambda_n)},X_1^{(\Lambda_{n+1})},X_2^{(\Lambda_{n+1})}\right)\right]\\
&+2P[N_{n+1}=1]P[N_n>1]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_1^{(\Lambda_{n+1})}\right),f\left(X_2^{(\Lambda_n)},X_1^{(\Lambda_{n+1})},X_1^{(\Lambda_{n+2})}\right)\right]\\
&+2P[N_{n+1}>1]P[N_n>1]\mathrm{Cov}\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_1^{(\Lambda_{n+1})}\right),f\left(X_2^{(\Lambda_n)},X_1^{(\Lambda_{n+1})},X_2^{(\Lambda_{n+1})}\right)\right]\\
&+V[N_n]\left\{E\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_3^{(\Lambda_n)}\right)\right]-\theta\right\}^2\\
&-2\theta P[N_n=1]\left\{E\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_3^{(\Lambda_n)}\right)\right]-E\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_1^{(\Lambda_{n+1})}\right)\right]\right\}\\
&+P[N_n=1]P[N_n>1]\left\{\left(E\left[f\left(X_1^{(\Lambda_n)},X_1^{(\Lambda_{n+1})},X_1^{(\Lambda_{n+2})}\right)\right]-E\left[f\left(X_1^{(\Lambda_n)},X_1^{(\Lambda_{n+1})},X_2^{(\Lambda_{n+1})}\right)\right]\right)^2\right.\\
&\left.-\left(E\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_3^{(\Lambda_n)}\right)\right]-E\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_1^{(\Lambda_{n+1})}\right)\right]\right)^2\right\}\\
&+2P[N_n=1]\left\{E\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_3^{(\Lambda_n)}\right)\right]-E\left[f\left(X_1^{(\Lambda_n)},X_2^{(\Lambda_n)},X_1^{(\Lambda_{n+1})}\right)\right]\right\}\\
&\quad\cdot\left\{P[N_n=1]E\left[f\left(X_1^{(\Lambda_n)},X_1^{(\Lambda_{n+1})},X_1^{(\Lambda_{n+2})}\right)\right]+P[N_n>1]E\left[f\left(X_1^{(\Lambda_n)},X_1^{(\Lambda_{n+1})},X_2^{(\Lambda_{n+1})}\right)\right]\right\}.
\end{aligned}
$$

In this case, there are two non-zero covariances:

$$
\begin{aligned}
&Cov\left[R_n^{f-\theta}, R_{n+1}^{f-\theta}\right] = \\
&= P[N_n > 1]P[N_{n+1} > 2]Cov\left[f\left(X_1^{(\Lambda_n)}, X_2^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}\right), f\left(X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}, X_3^{(\Lambda_{n+1})}\right)\right] \\
&\quad + P[N_n > 1]E\left[(N_n - 3)_+\right]Cov\left[f\left(X_1^{(\Lambda_n)}, X_2^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}\right), f\left(X_2^{(\Lambda_{n+1})}, X_3^{(\Lambda_{n+1})}, X_4^{(\Lambda_{n+1})}\right)\right] \\
&\quad + P[N_n > 1]P[N_{n+1} = 2]Cov\left[f\left(X_1^{(\Lambda_n)}, X_2^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}\right), f\left(X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}\right)\right] \\
&\quad + P[N_n > 1]P[N_{n+1} > 2]Cov\left[f\left(X_1^{(\Lambda_n)}, X_2^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}\right), f\left(X_2^{(\Lambda_{n+1})}, X_3^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}\right)\right] \\
&\quad + P[N_n > 1]P[N_{n+1} = 1]P[N_{n+2} = 1]Cov\left[f\left(X_1^{(\Lambda_n)}, X_2^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}\right), f\left(X_1^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}, X_1^{(\Lambda_{n+3})}\right)\right] \\
&\quad + P[N_n > 1]P[N_{n+1} > 1]P[N_{n+2} = 1]Cov\left[f\left(X_1^{(\Lambda_n)}, X_2^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}\right), f\left(X_2^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}, X_1^{(\Lambda_{n+3})}\right)\right] \\
&\quad + P[N_n > 1]P[N_{n+1} = 1]P[N_{n+2} > 1]Cov\left[f\left(X_1^{(\Lambda_n)}, X_2^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}\right), f\left(X_1^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}, X_2^{(\Lambda_{n+2})}\right)\right] \\
&\quad + P[N_n > 1]P[N_{n+1} > 1]P[N_{n+2} > 1]Cov\left[f\left(X_1^{(\Lambda_n)}, X_2^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}\right), f\left(X_2^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}, X_2^{(\Lambda_{n+2})}\right)\right] \\
&\quad + P[N_{n+1} > 2]Cov\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}\right), f\left(X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}, X_3^{(\Lambda_{n+1})}\right)\right] \\
&\quad + P[N_{n+1} > 3]Cov\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}\right), f\left(X_2^{(\Lambda_{n+1})}, X_3^{(\Lambda_{n+1})}, X_4^{(\Lambda_{n+1})}\right)\right] \\
&\quad + E\left[(N_n - 4)_+\right]Cov\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}\right), f\left(X_3^{(\Lambda_{n+1})}, X_4^{(\Lambda_{n+1})}, X_5^{(\Lambda_{n+1})}\right)\right] \\
&\quad + P[N_{n+1} = 2]Cov\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}\right), f\left(X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}\right)\right] \\
&\quad + P[N_{n+1} = 3]Cov\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}\right), f\left(X_2^{(\Lambda_{n+1})}, X_3^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}\right)\right] \\
&\quad + P[N_{n+1} > 3]Cov\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}\right), f\left(X_3^{(\Lambda_{n+1})}, X_4^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}\right)\right] \\
&\quad + P[N_{n+1} = 1]P[N_{n+2} = 1]Cov\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}\right), f\left(X_1^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}, X_1^{(\Lambda_{n+3})}\right)\right] \\
&\quad + P[N_{n+1} = 2]P[N_{n+2} = 1]Cov\left[f(X_1(\Lambda_n), X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}, f\left(X_2^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}, X_1^{(\Lambda_{n+3})}\right)\right] \\
&\quad + P[N_{n+1} > 2]P[N_{n+2} = 1]Cov\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}\right), f\left(X_3^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}, X_1^{(\Lambda_{n+3})}\right)\right] \\
&\quad + P[N_{n+1} = 1]P[N_{n+2} > 1]Cov\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}\right), f\left(X_1^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}, X_2^{(\Lambda_{n+2})}\right)\right] \\
&\quad + P[N_{n+1} = 2]P[N_{n+2} > 1]Cov\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}\right), f\left(X_2^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}, X_2^{(\Lambda_{n+2})}\right)\right] \\
&\quad + P[N_{n+1} > 2]P[N_{n+2} > 1]Cov\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}\right), f\left(X_3^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}, X_2^{(\Lambda_{n+2})}\right)\right] \\
&\quad + \theta P[N_n = 1]\left\{E\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}\right)\right] - E\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}\right)\right]\right\} \\
&\quad + \left\{P[N_n = 1]E\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}\right)\right] + P[N_n > 1]E\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}\right)\right]\right\}^2 \\
&\quad - E\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}\right)\right] \\
&\qquad \cdot\left\{P[N_n = 1]E\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_1^{(\Lambda_{n+2})}\right)\right] + P[N_n > 1]E\left[f\left(X_1^{(\Lambda_n)}, X_1^{(\Lambda_{n+1})}, X_2^{(\Lambda_{n+1})}\right)\right]\right\}
\end{aligned}
$$

$$
\begin{aligned}
&\mathrm{Cov}\left[\mathrm{R}_{\mathrm{n}}^{\mathrm{f}-\theta}, \mathrm{R}_{\mathrm{n}+2}^{\mathrm{f}-\theta}\right] = \mathrm{P}[\mathrm{N}_{\mathrm{n}+1}=1]\mathrm{P}[\mathrm{N}_{\mathrm{n}+2}>2] \\
&\times \mathrm{Cov}\left[\mathrm{f}\left(\mathrm{X}_1^{(\Lambda_{\mathrm{n}})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+1})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+2})}\right), \mathrm{f}\left(\mathrm{X}_1^{(\Lambda_{\mathrm{n}+2})}, \mathrm{X}_2^{(\Lambda_{\mathrm{n}+2})}, \mathrm{X}_3^{(\Lambda_{\mathrm{n}+2})}\right)\right] + \mathrm{P}[\mathrm{N}_{\mathrm{n}+1}=1]\mathrm{E}\left[(\mathrm{N}_{\mathrm{n}+2}-3)_+\right] \\
&\times \mathrm{Cov}\left[\mathrm{f}\left(\mathrm{X}_1^{(\Lambda_{\mathrm{n}})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+1})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+2})}\right), \mathrm{f}\left(\mathrm{X}_2^{(\Lambda_{\mathrm{n}+2})}, \mathrm{X}_3^{(\Lambda_{\mathrm{n}+2})}, \mathrm{X}_4^{(\Lambda_{\mathrm{n}+2})}\right)\right] + \mathrm{P}[\mathrm{N}_{\mathrm{n}+1}=1]\mathrm{P}[\mathrm{N}_{\mathrm{n}+2}=2] \\
&\times \mathrm{Cov}\left[\mathrm{f}\left(\mathrm{X}_1^{(\Lambda_{\mathrm{n}})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+1})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+2})}\right), \mathrm{f}\left(\mathrm{X}_1^{(\Lambda_{\mathrm{n}+2})}, \mathrm{X}_2^{(\Lambda_{\mathrm{n}+2})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+3})}\right)\right] + \mathrm{P}[\mathrm{N}_{\mathrm{n}+1}=1]\mathrm{P}[\mathrm{N}_{\mathrm{n}+2}>2] \\
&\times \mathrm{Cov}\left[\mathrm{f}\left(\mathrm{X}_1^{(\Lambda_{\mathrm{n}})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+1})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+2})}\right), \mathrm{f}\left(\mathrm{X}_2^{(\Lambda_{\mathrm{n}+2})}, \mathrm{X}_3^{(\Lambda_{\mathrm{n}+2})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+3})}\right)\right] \\
&\quad + \mathrm{P}[\mathrm{N}_{\mathrm{n}+1}=1]\mathrm{P}[\mathrm{N}_{\mathrm{n}+2}=1]\mathrm{P}[\mathrm{N}_{\mathrm{n}+3}=1] \\
&\times \mathrm{Cov}\left[\mathrm{f}\left(\mathrm{X}_1^{(\Lambda_{\mathrm{n}})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+1})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+2})}\right), \mathrm{f}\left(\mathrm{X}_1^{(\Lambda_{\mathrm{n}+2})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+3})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+4})}\right)\right] \\
&\quad + \mathrm{P}[\mathrm{N}_{\mathrm{n}+1}=1]\mathrm{P}[\mathrm{N}_{\mathrm{n}+2}=1]\mathrm{P}[\mathrm{N}_{\mathrm{n}+3}>1] \\
&\times \mathrm{Cov}\left[\mathrm{f}\left(\mathrm{X}_1^{(\Lambda_{\mathrm{n}})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+1})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+2})}\right), \mathrm{f}\left(\mathrm{X}_1^{(\Lambda_{\mathrm{n}+2})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+3})}, \mathrm{X}_2^{(\Lambda_{\mathrm{n}+3})}\right)\right] \\
&\quad + \mathrm{P}[\mathrm{N}_{\mathrm{n}+1}=1]\mathrm{P}[\mathrm{N}_{\mathrm{n}+2}>1]\mathrm{P}[\mathrm{N}_{\mathrm{n}+3}=1] \\
&\times \mathrm{Cov}\left[\mathrm{f}\left(\mathrm{X}_1^{(\Lambda_{\mathrm{n}})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+1})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+2})}\right), \mathrm{f}\left(\mathrm{X}_2^{(\Lambda_{\mathrm{n}+2})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+3})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+4})}\right)\right] \\
&\quad + \mathrm{P}[\mathrm{N}_{\mathrm{n}+1}=1]\mathrm{P}[\mathrm{N}_{\mathrm{n}+2}>1]\mathrm{P}[\mathrm{N}_{\mathrm{n}+3}>1] \\
&\times \mathrm{Cov}\left[\mathrm{f}\left(\mathrm{X}_1^{(\Lambda_{\mathrm{n}})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+1})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+2})}\right), \mathrm{f}\left(\mathrm{X}_2^{(\Lambda_{\mathrm{n}+2})}, \mathrm{X}_1^{(\Lambda_{\mathrm{n}+3})}, \mathrm{X}_2^{(\Lambda_{\mathrm{n}+3})}\right)\right].
\end{aligned}
$$

The preceding formulae can be used to compute the asymptotic variances $\gamma_{f_1}^2$ and $\gamma_{f_2}^2$ in Lemma 1; here is a sketchy derivation:

$$
\begin{aligned}
\gamma_{f_i}^2 &= \chi_0 + 2\chi_1 \\
&= \frac{1}{\alpha_1}\left\{ V\left[R_n^{f_i} - \theta_{f_i} N_n\right] + 2\mathrm{Cov}\left[R_n^{f_i} - \theta_{f_i} N_n, R_{n+1}^{f_i} - \theta_{f_i} N_{n+1}\right]\right\} \\
&= \frac{1}{\alpha_1}\left\{ (\alpha_2 - 2\alpha_1 + 1)m_4 + \left(2\alpha_2 - 7\alpha_1 - 2\pi + 6 + \frac{\alpha_2}{\alpha_1^2}\right)\tau^4 \right. \\
&\qquad + \left(2\alpha_1 - \alpha_2 - 1 + \frac{\alpha_2}{\alpha_1^2}\right)m_2^2 + 2\left(6\alpha_1 - 2\alpha_2 + \pi - 4 - \frac{\alpha_2}{\alpha_1^2}\right)m_2\tau^2 + \left(\frac{\alpha_2}{\alpha_1^2} - 2\right)m_1^4 \\
&\qquad + 2(\alpha_1 - 1)m_1 m_3 + 2\left(2 - \alpha_1 - \frac{\alpha_2}{\alpha_1^2}\right)m_2 m_1^2 + 2\left(-2\alpha_1 + 2 - \pi + \frac{\alpha_2}{\alpha_1^2}\right)m_1^2\tau^2 \\
&\qquad \left. + 2\left[- m_1^4 + (\alpha_1 - 1)m_1 m_3 + (2 - \alpha_1)m_2 m_1^2 + 2(1 - \alpha_1)m_1^2\tau^2\right]\right\} \\
&= \frac{1}{\alpha_1}\left\{ (\alpha_2 - 2\alpha_1 + 1)m_4 + \left(2\alpha_2 - 7\alpha_1 - 2\pi + 6 + \frac{\alpha_2}{\alpha_1^2}\right)\tau^4 + \left(2\alpha_1 - \alpha_2 - 1 + \frac{\alpha_2}{\alpha_1^2}\right)m_2^2 \right. \\
&\qquad + \left(12\alpha_1 - 4\alpha_2 + 2\pi - 8 - 2\frac{\alpha_2}{\alpha_1^2}\right)m_2\tau^2 + \left(\frac{\alpha_2}{\alpha_1^2} - 4\right)m_1^4 + 4(\alpha_1 - 1)m_1 m_3 \\
&\qquad \left. + 2\left(4 - 2\alpha_1 - \frac{\alpha_2}{\alpha_1^2}\right)m_2 m_1^2 - 2\left(4\alpha_1 + \pi - 4 - \frac{\alpha_2}{\alpha_1^2}\right)m_1^2\tau^2\right\};
\end{aligned}
$$

$$
\begin{aligned}
\gamma_{f_2}^2 &= \chi_0 + 2\chi_1 + 2\chi_2 \\
&= \frac{1}{\alpha_1}\Big\{V\big[R_n^{f_2} - \theta_{f_2}N_n\big] + 2\,\mathrm{Cov}\Big[R_n^{f_2} - \theta_{f_2}N_n, R_{n+1}^{f_2} - \theta_{f_2}N_{n+1}\Big] + 2\mathrm{Cov}\Big[R_n^{f_2} - \theta_{f_2}N_n, R_{n+2}^{f_2} - \theta_{f_2}N_{n+2}\Big]\Big\} \\
&= \frac{1}{\alpha_1}\Big\{(\alpha_2 - 4\alpha_1 + 4 - \pi)m_4 + \Big[2\alpha_2 - 11\alpha_1 - 2\big(\pi^3 - 5\pi^2 + 9\pi - 8\big) + 2\pi(2-\pi)\frac{1}{\alpha_1} + (2-\pi)^2\frac{\alpha_2}{\alpha_1^2}\Big]\tau^4 \\
&\quad + \Big[4\alpha_1 - \alpha_2 + 2\pi^2 - 3\pi - 4 + 2\pi(2-\pi)\frac{1}{\alpha_1} + (2-\pi)^2\frac{\alpha_2}{\alpha_1^2}\Big]m_2^2 \\
&\quad + \Big[20\alpha_1 - 4\alpha_2 + 2\big(\pi^3 - 6\pi^2 + 12\pi - 12\big) - 4\pi(2-\pi)\frac{1}{\alpha_1} - 2(2-\pi)^2\frac{\alpha_2}{\alpha_1^2}\Big]m_2\tau^2 \\
&\quad + \Big[4(\pi-2) + 2\pi(2-\pi)\frac{1}{\alpha_1} + (2-\pi)^2\frac{\alpha_2}{\alpha_1^2}\Big]m_1^4 + 4(\alpha_1 + \pi - 2)m_1m_3 \\
&\quad + \Big[-4\alpha_1 - 2(\pi-2)(\pi+4) - 4\pi(2-\pi)\frac{1}{\alpha_1} - 2(2-\pi)^2\frac{\alpha_2}{\alpha_1^2}\Big]m_2m_1^2 \\
&\quad + \Big[-8\alpha_1 + 2(2-\pi)\big(\pi^2 - 3\pi + 3\big) + 4\pi(2-\pi)\frac{1}{\alpha_1} + 2(2-\pi)^2\frac{\alpha_2}{\alpha_1^2}\Big]m_1^2\tau^2 \\
&\quad + 2\Big[-4(1-\pi)m_1^4 + \Big[\alpha_1(2-\pi) - (2-\pi)^2\Big]m_1m_3 + \big[\alpha_1(\pi-2) + \pi^2 - 8\pi + 8\big]m_2m_1^2 \\
&\quad + 2\big[\alpha_1(\pi-2) + \pi^2 - 3\pi + 3\big]m_1^2\tau^2\Big] + 2\Big[\pi(\pi-2)m_1^4 + [\pi\alpha_1 + \pi(\pi-2)]m_1m_3 \\
&\quad + [-\pi\alpha_1 + 2\pi(2-\pi)]m_2m_1^2 + [-2\pi\alpha_1 + \pi(3-\pi)]m_1^2\tau^2\Big]\Big\} \\
&= \frac{1}{\alpha_1}\Big\{(\alpha_2 - 4\alpha_1 + 4 - \pi)m_4 + \\
&\quad + \Big[2\alpha_2 - 11\alpha_1 - 2\big(\pi^3 - 5\pi^2 + 9\pi - 8\big) + (2-\pi)^2\frac{\alpha_2}{\alpha_1^2} + 2\pi(2-\pi)\frac{1}{\alpha_1}\Big]\tau^4 \\
&\quad + \Big[4\alpha_1 - \alpha_2 + 2\pi^2 - 3\pi - 4 + 2\pi(2-\pi)\frac{1}{\alpha_1} + (2-\pi)^2\frac{\alpha_2}{\alpha_1^2}\Big]m_2^2 \\
&\quad + 2\Big[10\alpha_1 - 2\alpha_2 + \pi^3 - 6\pi^2 + 12\pi - 12 - 2\pi(2-\pi)\frac{1}{\alpha_1} - (2-\pi)^2\frac{\alpha_2}{\alpha_1^2}\Big]m_2\tau^2 \\
&\quad + \Big[2\big(\pi^2 + 4\pi - 8\big) + 2\pi(2-\pi)\frac{1}{\alpha_1} + (2-\pi)^2\frac{\alpha_2}{\alpha_1^2}\Big]m_1^4 + 8[\alpha_1 + \pi - 2]m_1m_3 \\
&\quad - 2\Big[4\alpha_1 + 2\big(\pi^2 + 3\pi - 8\big) + 2\pi(2-\pi)\frac{1}{\alpha_1} + (2-\pi)^2\frac{\alpha_2}{\alpha_1^2}\Big]m_2m_1^2 \\
&\quad - 2\Big[8\alpha_1 + \pi^3 - 6\pi^2 + 12\pi - 12 - 2\pi(2-\pi)\frac{1}{\alpha_1} - (2-\pi)^2\frac{\alpha_2}{\alpha_1^2}\Big]m_1^2\tau^2\Big\}.
\end{aligned}
$$

Finally, we compute the explicit expression of the off-diagonal terms of $\mathbf{C}_{(\hat{\mathbf{m}}_1,\hat{\mathbf{m}}_2,\hat{\mathbf{m}}_{\mathbf{f}_1},\hat{\mathbf{m}}_{\mathbf{f}_2})}$.

$$
\begin{aligned}
\mathbf{C}_{(\hat{\mathbf{m}}_1,\hat{\mathbf{m}}_2,\hat{\mathbf{m}}_{\mathbf{f}_1},\hat{\mathbf{m}}_{\mathbf{f}_2})}(1,2) &= \mathbf{C}_{(\hat{\mathbf{m}}_1,\hat{\mathbf{m}}_2,\hat{\mathbf{m}}_{\mathbf{f}_1},\hat{\mathbf{m}}_{\mathbf{f}_2})}(2,1) = \mathrm{Cov}[\hat{m}_1,\hat{m}_2] \\
&= \frac{1}{\alpha_1}\mathrm{Cov}\Big[R_n^x - \theta_x N_n, R_n^{x^2} - \theta_{x^2}N_n\Big] \\
&= \frac{1}{\alpha_1}\big\{\alpha_2 m_3 - \alpha_2 m_1 m_2 + 2m_1\tau^2(\alpha_1 - \alpha_2)\big\}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{C}_{(\hat{\mathbf{m}}_1,\hat{\mathbf{m}}_2,\hat{\mathbf{m}}_{f_1},\hat{\mathbf{m}}_{f_2})}(1,3) &= \mathbf{C}_{(\hat{\mathbf{m}}_1,\hat{\mathbf{m}}_2,\hat{\mathbf{m}}_{f_1},\hat{\mathbf{m}}_{f_2})}(3,1) = \mathrm{Cov}\big[\hat{m}_1,\hat{m}_{f_1}\big] \\
&= \frac{1}{\alpha_1}\Big\{\mathrm{Cov}\big[R_n^x-\theta_x N_n, R_n^{f_1}-\theta_{f_1}N_n\big]+\mathrm{Cov}\big[R_{n+1}^x-\theta_x N_{n+1}, R_n^{f_1}-\theta_{f_1}N_n\big]\Big\} \\
&= \frac{1}{\alpha_1}\Big\{\Big[(\alpha_2-\alpha_1)m_3-\alpha_1 m_1^3+(2\alpha_1-\alpha_2)m_1m_2 \\
&\qquad +(3\alpha_1-2\alpha_2-1)m_1\tau^2\big]+\big[-\alpha_1 m_1^3+\alpha_1 m_1 m_2+(1-\alpha_1)m_1\tau^2\big]\Big\} \\
&= \frac{1}{\alpha_1}\Big\{(\alpha_2-\alpha_1)m_3-2\alpha_1 m_1^3+(3\alpha_1-\alpha_2)m_1m_2+2(\alpha_1-\alpha_2)m_1\tau^2\Big\}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{C}_{(\hat{\mathbf{m}}_1,\hat{\mathbf{m}}_2,\hat{\mathbf{m}}_{f_1},\hat{\mathbf{m}}_{f_2})}(2,3) &= \mathbf{C}_{(\hat{\mathbf{m}}_1,\hat{\mathbf{m}}_2,\hat{\mathbf{m}}_{f_1},\hat{\mathbf{m}}_{f_2})}(3,2) = \mathrm{Cov}\big[\hat{m}_2,\hat{m}_{f_1}\big] = \frac{1}{\alpha_1}\Big\{\mathrm{Cov}\Big[R_n^{x^2}-\theta_{x^2}N_n, R_n^{f_1}-\theta_{f_1}N_n\Big] \\
&\quad +\mathrm{Cov}\Big[R_{n+1}^{x^2}-\theta_{x^2}N_{n+1}, R_n^{f_1}-\theta_{f_1}N_n\Big]\Big\} \\
&= \frac{1}{\alpha_1}\Big\{\Big[(\alpha_2-\alpha_1)m_4+2(\alpha_2-3\alpha_1+2)\tau^4+(\alpha_1-\alpha_2)m_2^2 \\
&\quad +4(2\alpha_1-\alpha_2-1)m_2\tau^2+\alpha_1 m_1m_3-\alpha_1 m_2 m_1^2+2(1-\alpha_1)m_1^2\tau^2\Big] \\
&\quad +\Big[\alpha_i m_1m_3-\alpha_1 m_2m_1^2+2(1-\alpha_1)m_1^2\tau^2\Big]\Big\} \\
&= \frac{1}{\alpha_1}\Big\{(\alpha_2-\alpha_1)m_4+2(\alpha_2-3\alpha_1+2)\tau^4+(\alpha_1-\alpha_2)m_2^2 \\
&\quad +4(2\alpha_1-\alpha_2-1)m_2\tau^2 \quad +2\alpha_1 m_1m_3-2\alpha_1 m_2m_1^2+4(1-\alpha_1)m_1^2\tau^2\Big\}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{C}_{(\hat{\mathbf{m}}_1,\hat{\mathbf{m}}_2,\hat{\mathbf{m}}_{f_1},\hat{\mathbf{m}}_{f_2})}(1,4) &= \mathbf{C}_{(\hat{\mathbf{m}}_1,\hat{\mathbf{m}}_2,\hat{\mathbf{m}}_{f_1},\hat{\mathbf{m}}_{f_2})}(4,1) = \mathrm{Cov}\big[\hat{m}_1,\hat{m}_{f_2}\big] \\
&= \frac{1}{\alpha_1}\Big\{\mathrm{Cov}\big[R_n^x-\theta_x N_n, R_n^{f_2}-\theta_{f_2}N_n\big]+\mathrm{Cov}\big[R_{n+1}^x-\theta_x N_{n+1}, R_n^{f_2}-\theta_{f_2}N_n\big] \\
&\quad +\mathrm{Cov}\big[R_{n+2}^x-\theta_x N_{n+2}, R_n^{f_2}-\theta_{f_2}N_n\big]\Big\} \\
&= \frac{1}{\alpha_1}\Big\{\Big[(\alpha_2-2\alpha_1+\pi)m_3-(2\alpha_1-\pi)m_1^3+(4\alpha_1-\alpha_2-2\pi)m_1m_2 \\
&\quad +2(2\alpha_1-\alpha_2-1)m_1\tau^2\Big]+\Big[(\pi\alpha_1-2\alpha_1+\pi)m_1^3 \\
&\quad +(2\alpha_1-\pi\alpha_1-\pi)m_1m_2+(\pi\alpha_1-2\alpha_1+2-\pi)m_1\tau^2\Big] \\
&\quad +\Big[-\pi\alpha_1 m_1^3+\pi\alpha_1 m_1m_2+\pi(1-\alpha_1)m_1\tau^2\Big]\Big\} \\
&= \frac{1}{\alpha_1}\Big\{(\alpha_2-2\alpha_1+\pi)m_3-(4\alpha_1-2\pi)m_1^3+(6\alpha_1-\alpha_2-3\pi)m_1m_2 \\
&\quad +2(\alpha_1-\alpha_2)m_1\tau^2\Big\}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{C}_{(\hat{\mathbf{m}}_1,\hat{\mathbf{m}}_2,\hat{\mathbf{m}}_{f_1},\hat{\mathbf{m}}_{f_2})}(2,4) &= \mathbf{C}_{(\hat{\mathbf{m}}_1,\hat{\mathbf{m}}_2,\hat{\mathbf{m}}_{f_1},\hat{\mathbf{m}}_{f_2})}(4,2) = \mathrm{Cov}\big[\hat{m}_2,\hat{m}_{f_2}\big] \\
&= \frac{1}{\alpha_1}\Big\{\mathrm{Cov}\Big[R_n^{x^2}-\theta_{x^2}N_n, R_n^{f_2}-\theta_{f_2}N_n\Big]+\mathrm{Cov}\Big[R_{n+1}^{x^2}-\theta_{x^2}N_{n+1}, R_n^{f_2}-\theta_{f_2}N_n\Big] \\
&\quad +\mathrm{Cov}\Big[R_{n+2}^{x^2}-\theta_{x^2}N_{n+2}, R_n^{f_2}-\theta_{f_2}N_n\Big]\Big\} \\
&= \frac{1}{\alpha_1}\Big\{\Big[(\alpha_2-2\alpha_1+\pi)m_4+2(\alpha_2-4\alpha_1-\pi+4)\tau^4+(2\alpha_1-\alpha_2-\pi)m_2^2 \\
&\quad +4(3\alpha_1-\alpha_2-2)m_2\tau^2+(2\alpha_1-\pi)m_1m_3+(\pi-2\alpha_1)m_2m_1^2 \\
&\quad +4(1-\alpha_1)m_1^2\tau^2\Big]+\Big[(2\alpha_1-\pi\alpha_1-\pi)m_1m_3+(\pi\alpha_1-2\alpha_1+\pi)m_2m_1^2 \\
&\quad +2(\pi\alpha_1-2\alpha_1-\pi+2)m_1^2\tau^2\Big] \\
&\quad +\Big[\pi\alpha_1 m_1m_3-\pi\alpha_1 m_2m_1^2+2\pi(1-\alpha_1)m_1^2\tau^2\Big]\Big\} \\
&= \frac{1}{\alpha_1}\Big\{(\alpha_2-2\alpha_1+\pi)m_4+2(\alpha_2-4\alpha_1-\pi+4)\tau^4+(2\alpha_1-\alpha_2-\pi)m_2^2 \\
&\quad +4(3\alpha_1-\alpha_2-2)m_2\tau^2+2(2\alpha_1-\pi)m_1m_3+2(\pi-2\alpha_1)m_2m_1^2 \\
&\quad +8(1-\alpha_1)m_1^2\tau^2\Big\}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{C}_{(\hat{\mathbf{m}}_1,\hat{\mathbf{m}}_2,\hat{\mathbf{m}}_{\mathbf{f}_1},\hat{\mathbf{m}}_{\mathbf{f}_2})}(3,4) &= \mathbf{C}_{(\hat{\mathbf{m}}_1,\hat{\mathbf{m}}_2,\hat{\mathbf{m}}_{\mathbf{f}_1},\hat{\mathbf{m}}_{\mathbf{f}_2})}(4,3) = \mathrm{Cov}\left[\hat{m}_{f_1},\hat{m}_{f_2}\right] \\
&= \frac{1}{\alpha_1}\Bigg\{\mathrm{Cov}\left[R_n^{f_1}-\theta_{f_1}N_n, R_n^{f_2}-\theta_{f_2}N_n\right] + \mathrm{Cov}\left[R_{n+1}^{f_1}-\theta_{f_1}N_{n+1}, R_n^{f_2}-\theta_{f_2}N_n\right] \\
&\quad +\mathrm{Cov}\left[R_{n+2}^{f_1}-\theta_{f_1}N_{n+2}, R_n^{f_2}-\theta_{f_2}N_n\right] \\
&\quad +\mathrm{Cov}\left[R_{n-1}^{f_1}-\theta_{f_1}N_{n-1}, R_n^{f_2}-\theta_{f_2}N_n\right]\Bigg\} \\
&= \frac{1}{\alpha_1}\Bigg\{\Big[(\alpha_2-3\alpha_1+2)m_4 \\
&\quad +\left(2\alpha_2-10\alpha_1+2\pi^2-8\pi+12+\pi\frac{1}{\alpha_1}+(2-\pi)\frac{\alpha_2}{\alpha_1^2}\right)\tau^4 \\
&\quad +\left(3\alpha_1-\alpha_2-2\pi-2+\pi\frac{1}{\alpha_1}+(2-\pi)\frac{\alpha_2}{\alpha_1^2}\right)m_2^2 \\
&\quad +2\left(8\alpha_1-2\alpha_2-\pi^2+5\pi-8-\pi\frac{1}{\alpha_1}-(2-\pi)\frac{\alpha_2}{\alpha_1^2}\right)m_2\tau^2 \\
&\quad +\left(\pi-4+\pi\frac{1}{\alpha_1}+(2-\pi)\frac{\alpha_2}{\alpha_1^2}\right)m_1^4+(3\alpha_1+\pi-4)m_1m_3 \\
&\quad +\left(8-3\alpha_1-2\pi\frac{1}{\alpha_1}+2(\pi-2)\frac{\alpha_2}{\alpha_1^2}\right)m_2m_1^2 \\
&\quad +2\left(-3\alpha_1+\pi^2-4\pi+4+\pi\frac{1}{\alpha_1}+(2-\pi)\frac{\alpha_2}{\alpha_1^2}\right)m_1^2\tau^2\Big] \\
&\quad +\Big[(\pi-2)m_1^4+(\alpha_1-1)(2-\pi)m_1m_3+(2-\pi)(2-\alpha_1)m_2m_1^2 \\
&\quad +[2\alpha_1(\pi-2)+5-3\pi]m_1^2\tau^2\Big] \\
&\quad +\Big[-\pi m_1^4+\pi(\alpha_1-1)m_1m_3+\pi(2-\alpha_1)m_2m_1^2+2\pi(1-\alpha_1)m_1^2\tau^2\Big] \\
&\quad +\Big[(\pi-2)m_1^4+(\alpha_1+\pi-2)m_1m_3-(\alpha_1+2\pi-4)m_2m_1^2 \\
&\quad +(-2\alpha_1-\pi+3)m_1^2\tau^2\Big]\Bigg\} \\
&= \frac{1}{\alpha_1}\Bigg\{(\alpha_2-3\alpha_1+2)m_4 \\
&\quad +\left[2\alpha_2-10\alpha_1+2\pi^2-8\pi+12+\pi\frac{1}{\alpha_1}+(2-\pi)\frac{\alpha_2}{\alpha_1^2}\right]\tau^4 \\
&\quad +\left[3\alpha_1-\alpha_2-2\pi-2+\pi\frac{1}{\alpha_1}+(2-\pi)\frac{\alpha_2}{\alpha_1^2}\right]m_2^2 \\
&\quad +2\left[8\alpha_1-2\alpha_2-\pi^2+5\pi-8-\pi\frac{1}{\alpha_1}-(2-\pi)\frac{\alpha_2}{\alpha_1^2}\right]m_2\tau^2 \\
&\quad +\left[2\pi-8+\pi\frac{1}{\alpha_1}+(2-\pi)\frac{\alpha_2}{\alpha_1^2}\right]m_1^4 \\
&\quad +2(3\alpha_1+\pi-4)m_1m_3-2\left[3\alpha_1+\pi-8+\pi\frac{1}{\alpha_1}+(2-\pi)\frac{\alpha_2}{\alpha_1^2}\right]m_2m_1^2 \\
&\quad +2\left[-6\alpha_1+\pi^2-5\pi+8+\pi\frac{1}{\alpha_1}+(2-\pi)\frac{\alpha_2}{\alpha_1^2}\right]m_1^2\tau^2\Bigg\}.
\end{aligned}
$$

Appendix B. Correcting Cline's error

CLINE's paper (1983) assumes normality of the level distribution and thus derives formulae under that assumption. Such formulae can be derived by those computed here, in particular from those expressed only in terms of π, which appears also in CLINE (1983), τ, appearing in CLINE (1983) with a different parametrization, and m_i's, which are the moments of the level distribution and can thus be derived in terms of Cline's parameters. The needed parameter change is thus

$$\begin{aligned}
\pi &= \pi \\
\tau^2 &= (1-\rho)\sigma^2 \\
m_1 &= \boldsymbol{\mu}_1 = \boldsymbol{\mu} \\
m_2 &= \boldsymbol{\mu}_2 + \tau^2 = \mu^2 + \sigma^2 \\
m_3 &= \mu_3 + 3\mu_1\tau^2 = \mu^3 + 3\mu\sigma^2 \\
m_4 &= \mu_4 + 6\mu_2\tau^2 + 3\tau^4 = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4.
\end{aligned}$$

We thus checked all of Cline's expressions, finding two errors, which we report to make Cline's formulae directly usable.

The first error concerns $\mathrm{Cov}\left[R_n^x - \theta_x N_n, R_n^{f_1} - \theta_{f_1} N_n\right]$, and it is simply a typo because the subsequent formulae use the correct expression:

$$\begin{aligned}
\mathrm{Cov}\left[R_n^x - \theta_x N_n, R_n^{f_1} - \theta_{f_1} N_n\right] &= (\alpha_2 - \alpha_1)m_3 - \alpha_1 m_1^3 + (2\alpha_1 - \alpha_2)m_1 m_2 \\
&\quad + (3\alpha_1 - 2\alpha_2 - 1)m_1\tau^2 \\
&= \frac{1}{\pi}\left[(2-\pi) + \rho\frac{(1-\pi)(4-\pi)}{\pi}\right]\mu\sigma^2
\end{aligned}$$

where the first equality follows from the GFN model , whereas the second is the one with GNN parameters.

Instead, the second error is not directly comparable with one of our asymptotic value as it appears in calculation of asymptotic distribution of the second autocovariances present in Cline but not in our model. However, it can be retrieved by calculating the asymptotic variance of the autocovariance γ_2, as it is denoted in Cline paper, through a multidimensional delta method characterized by the following elements: the function

$$\begin{array}{rccc}
\gamma_2: & \mathbb{R}^2 & \to & \mathbb{R} \\
 & \left(\hat{m}_1, \hat{m}_{f_2}\right) & \mapsto & \gamma_2 = \hat{m}_{f_2} - \hat{m}_1^2
\end{array}$$

whose gradient is $\nabla\gamma_2\left(m_1, , m_{f_2}\right) = (-2m_1, 1)$ and the variance–covariance matrix

$$\mathbf{C}_{(\hat{m}_1, \hat{m}_{f_2})} = \begin{bmatrix} \gamma_1^2 & \mathrm{Cov}\left[\hat{m}_1, , \hat{m}_{f_2}\right] \\ \mathrm{Cov}\left[\hat{m}_1, , \hat{m}_{f_2}\right] & \gamma_{f_2}^2 \end{bmatrix}.$$

where all elements are already known. Then it follows that

$$V[\gamma_2] = \frac{1}{n}\nabla\gamma_2 \mathbf{C}_{(\hat{m}_{1,},\hat{m}_{f_2})}\nabla\gamma_2^t = \frac{\sigma^4}{n}\left[1 + (1-\pi)^4\left(2\rho - 5\rho^2\right) + 4\frac{(1-\pi)^2}{\pi}\rho^2\right]$$

where the second equality is obtained by replacing the GFN-SLM parameters with the GNN-SLM ones.

We finally verify that the formula in CLINE (1983) is incorrect. This can be easily observed in the case $\pi = 1$; this is a perfectly acceptable range of parameters for the GNN-SLM model, whereas our derivation, albeit carried out on the assumption that $\pi < 1$, does not actually depend on that assumption for $V[\gamma_1]$ and $V[\gamma_2]$. In such case, the X_i's are independent, and thus, there should be no difference between the two autocovariances defined by Cline, γ_1 and γ_2. As a consequence, the asymptotic distributions of $\hat{\gamma}_1$ and $\hat{\gamma}_2$ should be the same, and in particular, the two asymptotic variances should be the same, that is, $V[\gamma_1] = V[\gamma_2]$. We report in the following the two expressions as they appear in Cline:

$$V[\gamma_1] = \frac{\sigma^4}{n}\left[1 + \left(2\rho - 3\rho^2\right)(1-\pi)^2 + \frac{2(1-\pi)(2-\pi)}{\pi}\rho^2\right]$$

$$V[\gamma_2] = \frac{\sigma^4}{n}\left\{1 + \left(2\rho - 5\rho^2\right)(1-\pi)^4 + 2\rho^2\left[\frac{(1-\pi)^2 + (2-\pi)^2 - (1-\pi)^4}{\pi(2-\pi)}\right]\right\}$$

If we evaluate the latter expression for $\pi = 1$, we obtain that $V[\gamma_2] = (1 + 2\rho^2)\frac{\sigma^4}{n}$, instead of the value $V[\gamma_1] = \frac{\sigma^4}{n}$, which coincides with that of the variance $V[\gamma_2]$ that we have calculated earlier.

REFERENCES

BAYANI, J., S. SELVARAJAH, G. MAIRE, B. VUKOVIC, K. AL-ROMAIH, M. ZIELENSKA and J. A. SQUIRE (2007), Genomic mechanisms and measurement of structural and numerical instability in cancer cells, *Seminars in Cancer Biology* **17**, 5–18

BENELLI, M., G. MARSEGLIA, G. NANNETTI, R. PARAVIDINO, F. ZARA, F. D. BRICARELLI, F. TORRICELLI and A. MAGI (2010), A very fast and accurate method for calling aberrations in array-CGH data, *Biostatistics* **11**, 515–518

BREDEL, M., C. BREDEL, D. JURIC, G. R. HARSH, H. VOGEL, L. D. RECHT and B. I. SIKIC (2005), High-resolution genomic-wide mapping of genetic alterations in human glial brain tumors, *Cancer Research* **65**, 4088–4096.

CARTER N. P. (2007), Methods and strategies for analyzing copy number variation using DNA microarrays, *Nature Genetics* **39**, S16–S21.

CHERNOFF, H. and S. ZACKS (1964), Estimating the current mean of a normal distribution which is subjected to change in time. *The Annals of Mathematical Statistics* **35**, 999–1018.

CLINE, D. B. H. (1983), Limit theorems for the shifting level process, *Journal of Applied probability* **20**, 322–337.

FORNEY, G. D. (1973), The Viterbi algorithm, *Proceedings of the IEEE* **61**, 268–278.

FORTIN, V. and A. KEHAGIAS (2006), Time series segmentation with shifting means hidden Markov models. *Nonlinear Processes in Geophysics* **13**, 135–163.

FRIDLYAND, J., A. M. SNIJDERS, D. PINKEL, D. G. A. ALBERTSON and A. N. JAIN (2004), Hidden Markov models approach to the analysis of array-CGH data, *Journal of Multivariate Analysis* **90**, 132–153.

HUPE P., N. STRANSKY, J. P. THIERY, F. RADVANYI and E. BARILLOT (2004), Analysis of array CGH data: from signal ratio to gain and loss of DNA regions, *Bioinformatics* **20**, 3413–3422.

LAI, W. R. R., M. D. D. JOHNSON, R. KUCHERLAPATI and P. J. J. PARK (2005), Comparative analysis of algorithms for identifying amplifications and deletions in array-CGH data, *Bioinformatics* **21**, 3763–3770.

LEPRETRE, F., C. VILLENET, S. QUIEF, O. NIBOUREL, C. JACQUEMIN, X. TROUSSARD, F. JARDIN, F. GIBSON, J. P. KERCKAERT, C. ROUMIER and M. FIGEAC (2010), Waved aCGH: to smooth or not to smooth, *Nucleic Acids Research* **38**, e94.

LIU, X. S. (2007), Getting started in tiling microarray analysis, *PLoS Computational Biology* **3**, e183, 1842–1844.

MAGI, A., M. BENELLI, G. MARSEGLIA, G. NANNETTI, M. R. SCORDO and F. TORRICELLI (2010), A shifting level model algorithm that identifies aberrations in array-CGH data. *Biostatistics* **11**(2), 265.

MYERS C. L., M. J. DUNHAM, S. Y. KUNG and O. G. TROYANSKAYA (2004), Accurate detection of aneuploidies in array CGH and gene expression microarray data, *Bioinformatics* **20**, 3533–3543.

OLSHEN, A. B., E. S. VENKATRAMAN, R. LUCITO and M. WIGLER (2004), Circular binary segmentation for the analysis of array-based DNA copy number data, *Biostatistics* **5**, 557–72.

OOSTLANDER, A. E., G. A. MEIJER and B. YLSTRA (2004), Microarray-based comparative genomic hybridization and its applications in human genetics, *Clinical Genetics* **66**, 488–495.

PICARD, F., S. ROBIN, M. LAVIELLE, C. VAISSE and J.-J. DAUDIN (2005), A statistical approach for array CGH data analysis, *BMC Bioinformatics* **6**, 1–14.

SALAS, J. D. and D. C. BOES (1980), Shifting level modelling of hydrologic time series, *Advances in Water Resources* **3**, 59–63.

VAN DE WIEL, M., I. K. KIM, S. J. VOSSE, W. N. VAN WIERINGEN, S. M. WILTING and B. YLSTRA (2007), CGHcall: calling aberrations for array CGH tumor profiles, *Bioinformatics* **7**, 892–894.

YOON, S., Z. XUAN, V. MAKAROV, K. YE and J. SEBAT (2009), Sensitive and accurate detection of copy number variants using read depth of coverage, *Genome Research* **19**, 1586–1592.

Received: 30 January 2011. Revised: 31 January 2013.